***BIOSTATISTICS\_Daniel\_C06***

p178

CHAPTER OVERVIEW

This chapter covers estimation, one of the two types of statistical inference. As discussed in earlier chapters, statistics, such as means and variances, can be calculated from samples drawn from populations. These statistics serve as estimates of the corresponding population parameters. We expect these estimates to differ by some amount from the parameters they estimate. This chapter introduces estimation procedures that take these differences into ac-

count, thereby providing a foundation for statistical inference procedures discussed in the remaining chapters of the book.

Este capítulo cubre la **estimación**, uno de los dos tipos de **inferencia estadística**. Como se analizó en capítulos anteriores, las **estadísticas**, como las medias y las varianzas, se pueden calcular a partir de muestras extraídas de poblaciones. Estas estadísticas sirven como estimaciones de los **parámetros poblacionales** correspondientes. Esperamos que estas estimaciones **difieran en cierta medida** de los parámetros que estiman. Este capítulo presenta **procedimientos de estimación** que tienen en cuenta estas diferencias, proporcionando así una base para los procedimientos de inferencia estadística que se analizan en los capítulos restantes del libro.

TOPICS

6.1 INTRODUCTION

6.2 CONFIDENCE INTERVAL FOR A POPULATION MEAN

6.3 THE t DISTRIBUTION

6.4 CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN TWO POPULATION

MEANS

6.5 CONFIDENCE INTERVAL FOR A POPULATION PROPORTION

6.6 CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN TWO POPULATION

PROPORTIONS

6.7 DETERMINATION OF SAMPLE SIZE FOR ESTIMATING MEANS

6.8 DETERMINATION OF SAMPLE SIZE FOR ESTIMATING PROPORTIONS

6.9 CONFIDENCE INTERVAL FOR THE VARIANCE OF A NORMALLY DISTRIBUTED

POPULATION

6.10 CONFIDENCE INTERVAL FOR THE RATIO OF THE VARIANCES OF TWO

NORMALLY DISTRIBUTED POPULATIONS

6.11 SUMMARY

LEARNING OUTCOMES

After studying this chapter, the student will

1. understand the importance and basic principles of estimation.

Comprender la importancia y los principios básicos de la estimación.

2. be able to calculate interval estimates for a variety of parameters.

Ser capaz de calcular estimaciones de intervalo para una variedad de parámetros.

3. be able to interpret a confidence interval from both a practical and a probabilistic viewpoint.

Ser capaz de interpretar un intervalo de confianza desde un punto de vista tanto práctico como probabilístico.

4. understand the basic properties and uses of the t distribution, chi-square distribution, and F distribution.

comprender las propiedades básicas y los usos de la distribución t, la distribución chi-cuadrado y la distribución F.

6.1 INTRODUCTION

We come now to a consideration of estimation, the first of the two general areas of statistical inference. The second general area, hypothesis testing, is examined in the next chapter. We learned in Chapter 1 that inferential statistics is defined as follows.

Pasamos ahora a considerar la estimación, la primera de las dos áreas generales de la inferencia estadística. La segunda área general, la prueba de hipótesis, se examina en el próximo capítulo. En el capítulo 1 aprendimos que la estadística inferencial se define de la siguiente manera.

DEFINITION

Statistical inference is the procedure by which we reach a conclusion about a population on the basis of the information contained in a sample drawn from that population.

La inferencia estadística es el procedimiento mediante el cual llegamos a una conclusión sobre una población a partir de la información contenida en una muestra extraída de esa población.

The process of estimation entails calculating, from the data of a sample, some statistic that is offered as an approximation of the corresponding parameter of the population from which the sample was drawn.

El proceso de estimación implica calcular, a partir de los datos de una muestra, algún estadístico que se ofrece como una aproximación del parámetro correspondiente de la población de la que se extrajo la muestra.

The rationale behind estimation in the health sciences field rests on the assumption that workers in this field have an interest in the parameters, such as means and proportions, of various populations. If this is the case, there is a good reason why one must rely on estimating procedures to obtain information regarding these parameters. Many populations of interest, although finite, are so large that a 100 percent examination would be prohibitive from the standpoint of cost.

La lógica detrás de la estimación en el campo de las ciencias de la salud se basa en el supuesto de que los trabajadores en este campo tienen interés en los parámetros, como las medias y las proporciones, de diversas poblaciones. Si este es el caso, hay una buena razón por la cual uno debe confiar en procedimientos de estimación para obtener información sobre estos parámetros. Muchas poblaciones de interés, aunque finitas, son tan grandes que un examen del 100 por ciento sería prohibitivo desde el punto de vista del costo.

Suppose the administrator of a large hospital is interested in the mean age of patients admitted to his hospital during a given year. He may consider it too expensive to go through the records of all patients admitted during that particular year and, consequently, elect to examine a sample of the records from which he can compute an estimate of the mean age of patients admitted that year.

Supongamos que el administrador de un hospital grande está interesado en la edad media de los pacientes ingresados en su hospital durante un año determinado. Puede considerar demasiado costoso revisar los registros de todos los pacientes admitidos durante ese año en particular y, en consecuencia, optar por examinar una muestra de los registros a partir de la cual pueda calcular una estimación de la edad media de los pacientes admitidos ese año.

A physician in general practice may be interested in knowing what proportion of a certain type of individual, treated with a particular drug, suffers undesirable side effects. No doubt, her concept of the population consists of all those persons who ever have been or ever will be treated with this drug. Deferring a conclusion until the entire population has been observed could have an adverse effect on her practice.

Un médico de práctica general puede estar interesado en saber qué proporción de un determinado tipo de individuo, tratado con un fármaco concreto, sufre **efectos secundarios** indeseables. Sin duda, su concepto de población consiste en todas aquellas personas que alguna vez han sido o serán tratadas con este medicamento. Aplazar una conclusión hasta que se haya observado a toda la población podría tener un efecto adverso en su práctica.

These two examples have implied an interest in estimating, respectively, a population mean and a population proportion. Other parameters, the estimation of which we will cover in this chapter, are the difference between two means, the difference between two proportions, the population variance, and the ratio of two variances.

Estos dos ejemplos han implicado un interés en estimar, respectivamente, una media poblacional y una proporción poblacional. Otros parámetros, cuya estimación cubriremos en este capítulo, son la diferencia entre dos medias, la diferencia entre dos proporciones, la varianza poblacional y la relación de dos varianzas.

We will find that for each of the parameters we discuss, we can compute two types of estimate: a point estimate and an interval estimate.

Encontraremos que para cada uno de los parámetros que analizamos, podemos calcular dos tipos de estimación: una estimación puntual y una estimación de intervalo.

DEFINITION

A point estimate is a single numerical value used to estimate the corresponding population parameter.

Una estimación puntual es un valor numérico único que se utiliza para estimar el parámetro poblacional correspondiente.

DEFINITION

An interval estimate consists of two numerical values defining a range of values that, with a specified degree of confidence, most likely includes the parameter being estimated.

Una estimación de intervalo consta de dos valores numéricos que definen un rango de valores que, con un **grado específico de confianza**, probablemente incluya el parámetro que se está estimando.

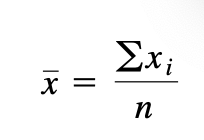
These concepts will be elaborated on in the succeeding sections.

Estos conceptos se desarrollarán en las secciones siguientes.

**Choosing an Appropriate Estimator**

Note that a single computed value has been referred to as an estimate. The rule that tells us how to compute this value, or estimate, is referred to as an estimator. Estimators are usually presented as formulas. For example,

Tenga en cuenta que a un único valor calculado se le denomina estimación. La regla que nos dice cómo calcular este valor, o estimación, se denomina **estimador**. Los estimadores suelen presentarse como fórmulas. Por ejemplo,



is an estimator of the population mean, mu. The single numerical value that results from evaluating this formula is called an estimate of the parameter mu.

In many cases, a parameter may be estimated by more than one estimator. For example, we could use the sample median to estimate the population mean. How then do we decide which estimator to use for estimating a given parameter? The decision is based on an objective measure or set of criteria that reflect some desired property of a particular estimator. When measured against these criteria, some estimators are better than others. One of these criteria is the property of unbiasedness.

es un estimador de la media poblacional, mu. El valor numérico único que resulta de evaluar esta fórmula se llama estimación del parámetro mu.

En muchos casos, un parámetro puede ser **estimado por más de un estimador**. Por ejemplo, podríamos usar la mediana muestral para estimar la media poblacional. ¿Cómo decidimos entonces qué estimador utilizar para estimar un parámetro determinado? La decisión se basa en una **medida objetiva** o un conjunto de criterios que reflejan alguna propiedad deseada de un estimador particular. Cuando se comparan con estos criterios, algunos estimadores son mejores que otros. Uno de estos criterios es la **propiedad de imparcialidad.**

**DEFINITION**

An estimator, say, T, of the parameter teta is said to be an unbiased estimator of teta if E(T) = teta

E(T) is read, “the expected value of T.” For a finite population, E(T) is obtained by taking the average value of T computed from all possible samples of a given size that may be drawn from the population. That is E(T) = muT. For an infinite population, E(T) is defined in terms of calculus.

E(T) se lee, “el valor esperado de T”. Para una población finita, E(T) se obtiene tomando el valor promedio de T calculado a partir de todas las muestras posibles de un tamaño determinado que puedan extraerse de la población. Esto es E(T) = muT. Para una población infinita, E(T) se define en términos de cálculo.

In the previous chapter we have seen that the sample mean, the sample proportion, the difference between two sample means, and the difference between two sample proportions are each unbiased estimates of their corresponding parameters. This property was implied when the parameters were said to be the means of the respective sampling distributions. For example, since the mean of the sampling distribution of is equal xm is equal to mu we know that is an unbiased estimator of The other criteria of good estimators will not be discussed in this book. The interested reader will find them covered in detail in most mathematical statistics texts.

En el capítulo anterior vimos que la **media muestral**, la proporción muestral, la diferencia entre dos medias muestrales y la diferencia entre dos proporciones muestrales son estimaciones **insesgadas** de sus parámetros correspondientes. Esta propiedad se implicó cuando se dijo que los parámetros eran las medias de las respectivas distribuciones muestrales. Por ejemplo, dado que la media de la distribución muestral de es igual a xm es igual a mu, sabemos que es un estimador insesgado de . Los demás criterios de buenos estimadores no se analizarán en este libro. El lector interesado los encontrará cubiertos en detalle en la mayoría de los textos de estadística matemática.

**Sampled Populations and Target Populations**

The health researcher who uses statistical inference procedures must be aware of the difference between two kinds of population—the sampled population and the target population.

El investigador de salud que utiliza procedimientos de inferencia estadística debe ser consciente de la diferencia entre dos tipos de población: la población muestreada y la población objetivo.

**DEFINITION**

The sampled population is the population from which one actually draws a sample.

La **población muestreada** es la población de la que realmente se extrae una muestra.

**DEFINITION**

The target population is the population about which one wishes to make an inference.

La **población objetivo** es la población sobre la cual se desea hacer una inferencia.

These two populations may or may not be the same. Statistical inference procedures allow one to make inferences about sampled populations (provided proper sampling methods have been employed). Only when the target population and the sampled population are the same is it possible for one to use statistical inference procedures to reach conclusions about the target population. If the sampled population and the target population are different, the researcher can reach conclusions about the target population only on the basis of nonstatistical considerations.

Estas dos poblaciones pueden ser iguales o no. Los procedimientos de inferencia estadística permiten hacer inferencias sobre poblaciones muestreadas (siempre que se hayan empleado métodos de muestreo adecuados). Sólo cuando la población objetivo y la población muestreada son las mismas es posible utilizar procedimientos de inferencia estadística para llegar a conclusiones sobre la población objetivo. **Si la población muestreada y la población objetivo son diferentes, el investigador puede llegar a conclusiones sobre la población objetivo sólo sobre la base de consideraciones no estadísticas.**

Suppose, for example, that a researcher wishes to assess the effectiveness of some method for treating rheumatoid arthritis. The target population consists of all patients suffering from the disease. It is not practical to draw a sample from this population. The researcher may, however, select a sample from all rheumatoid arthritis patients seen in some specific clinic. These patients constitute the sampled population, and, if proper sampling methods are used, inferences about this sampled population may be drawn on the basis of the information in the sample. If the researcher wishes to make inferences about all rheumatoid arthritis sufferers, he or she must rely on nonstatistical means to do so. Perhaps the researcher knows that the sampled population is similar, with respect to all important characteristics, to the target population. That is, the researcher may know that the age, sex, severity of illness, duration of illness, and so on are similar in both populations. And on the strength of this knowledge, the researcher may be willing to extrapolate his or her findings to the target population.

Supongamos, por ejemplo, que un investigador desea evaluar la eficacia de algún método para tratar la **artritis reumatoide**. La **población objetivo** está formada por todos los pacientes que padecen la enfermedad. No es práctico extraer una muestra de esta población. Sin embargo, el investigador puede seleccionar una muestra de todos los pacientes con artritis reumatoide atendidos en alguna clínica específica. Estos pacientes constituyen la población de la muestra y, si se utilizan métodos de muestreo adecuados, se pueden hacer inferencias sobre esta población de la muestra sobre la base de la información de la muestra. Si el investigador desea hacer inferencias sobre todos los que padecen artritis reumatoide, **debe confiar en medios no estadísticos para hacerlo**. Quizás el investigador sepa que la población muestreada es similar, con respecto a todas las características importantes, a la población objetivo. Es decir, el investigador puede saber que la edad, el sexo, la gravedad de la enfermedad, la duración de la enfermedad, etc., son similares en ambas poblaciones. Y basándose en este conocimiento, el investigador puede estar dispuesto a extrapolar sus hallazgos a la población objetivo.

In many situations the sampled population and the target population are identical; when this is the case, inferences about the target population are straightforward. The researcher, however, should be aware that this is not always the case and not fall into the trap of drawing unwarranted inferences about a population that is different from the one that is sampled.

En muchas situaciones, la población muestreada y la población objetivo son idénticas; cuando este es el caso, las inferencias sobre la población objetivo son sencillas. Sin embargo, el investigador debe ser consciente de que este no es siempre el caso y no caer en la trampa de hacer inferencias injustificadas sobre una población diferente de la que se muestra.

**Random and Nonrandom Samples**

In the examples and exercises of this book, we assume that the data available for analysis have come from random samples. The strict validity of the statistical procedures discussed depends on this assumption. In many instances in real-world applications it is impossible or impractical to use truly random samples. In animal experiments, for example, researchers usually use whatever animals are available from suppliers or their own breeding stock. If the researchers had to depend on randomly selected material, very little research of this type would be conducted. Again, nonstatistical considerations must play a part in the generalization process. Researchers may contend that the samples actually used are equivalent to simple random samples, since there is no reason to believe that the material actually used is not representative of the population about which inferences are desired.

En los ejemplos y ejercicios de este libro, asumimos que los datos disponibles para el análisis provienen de muestras aleatorias. La estricta validez de los procedimientos estadísticos discutidos depende de este supuesto. En muchos casos, en aplicaciones del mundo real, es imposible o poco práctico utilizar muestras verdaderamente aleatorias. En los experimentos con animales, por ejemplo, los investigadores suelen utilizar cualquier animal disponible de los proveedores o de sus propios reproductores. Si los investigadores tuvieran que depender de material seleccionado al azar, se realizarían muy pocas investigaciones de este tipo. Una vez más, las consideraciones no estadísticas deben desempeñar un papel en el proceso de generalización. Los investigadores pueden sostener que las muestras realmente utilizadas son equivalentes a muestras aleatorias simples, ya que no hay razón para creer que el material realmente utilizado no sea representativo de la población sobre la cual se desean hacer inferencias.

In many health research projects, samples of convenience, rather than random samples, are employed. Researchers may have to rely on volunteer subjects or on readily available subjects such as students in their classes. Samples obtained from such sources are examples of convenience samples. Again, generalizations must be made on the basis of nonstatistical considerations. The consequences of such generalizations, however, may be useful or they may range from misleading to disastrous.

En muchos proyectos de investigación en salud se emplean muestras de conveniencia, en lugar de muestras aleatorias. Es posible que los investigadores tengan que depender de sujetos voluntarios o de sujetos fácilmente disponibles, como los estudiantes de sus clases. Las muestras obtenidas de dichas fuentes son ejemplos de muestras de conveniencia. Nuevamente, las generalizaciones deben hacerse sobre la base de consideraciones no estadísticas. Sin embargo, las consecuencias de tales generalizaciones pueden ser útiles o pueden variar desde engañosas hasta desastrosas.

In some situations it is possible to introduce randomization into an experiment even though available subjects are not randomly selected from some well-defined population. In comparing two treatments, for example, each subject may be randomly assigned to one or the other of the treatments. Inferences in such cases apply to the treatments and not the subjects, and hence the inferences are valid.

En algunas situaciones es posible introducir la aleatorización en un experimento aunque los sujetos disponibles no se seleccionen al azar de una población bien definida. Al comparar dos tratamientos, por ejemplo, cada sujeto puede ser asignado aleatoriamente a uno u otro de los tratamientos. Las inferencias en tales casos se aplican a los tratamientos y no a los sujetos y, por tanto, las inferencias son válidas.

6.2 CONFIDENCE INTERVAL FOR A POPULATION MEAN

Suppose researchers wish to estimate the mean of some normally distributed population. They draw a random sample of size n from the population and compute xm, which they use as a point estimate of mu. Although this estimator of mu possesses all the qualities of a good estimator, we know that because random sampling inherently involves chance, xm cannot be expected to be equal to mu.

Supongamos que los investigadores desean estimar la media de alguna población distribuida normalmente. Extraen una muestra aleatoria de tamaño n de la población y calculan xm, que utilizan como estimación puntual de mu. Aunque este estimador de mu posee todas las cualidades de un buen estimador, sabemos que debido a que el muestreo aleatorio implica inherentemente azar, no se puede esperar que xm sea igual a mu.

It would be much more meaningful, therefore, to estimate mu by an interval that somehow communicates information regarding the probable magnitude of mu.

Por lo tanto, sería mucho más significativo estimar mu mediante un intervalo que de alguna manera comunique información sobre la magnitud probable de mu.

**Sampling Distributions and Estimation**

To obtain an interval estimate, we must draw on our knowledge of sampling distributions. In the present case, because we are concerned with the sample mean as an estimator of a population mean, we must recall what we know about the sampling distribution of the sample mean.

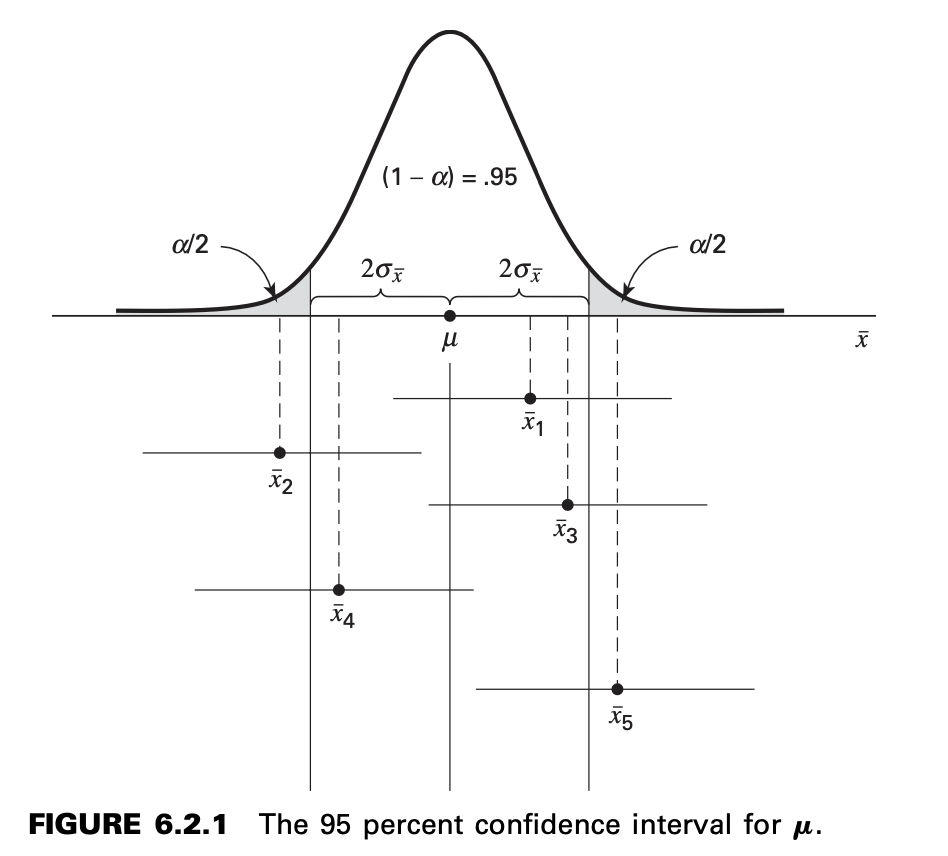
Para obtener una estimación de intervalo, debemos aprovechar nuestro conocimiento de las distribuciones muestrales. En el presente caso, como nos interesa la media muestral como estimador de una media poblacional, debemos recordar lo que sabemos sobre la distribución muestral de la media muestral.

In the previous chapter we learned that if sampling is from a normally distributed population, the sampling distribution of the sample mean will be normally distributed with a mean mux equal to the population mean mu, and a variance sigx equal to sig/n. We could plot the sampling distribution if we only knew where to locate it on the xm-axis.

En el capítulo anterior, aprendimos que si el muestreo se realiza a partir de una población con distribución normal, la distribución muestral de la media muestral se distribuirá normalmente, con una media mux igual a la media poblacional mu y una varianza sigx igual a sig/n. Podríamos representar gráficamente la distribución muestral si supiéramos dónde ubicarla en el eje xm.

From our knowledge of normal distributions, in general, we know even more about the distribution of in this case. We know, for example, that regardless of where the distribution of is located, approximately 95 percent of the possible values of constituting the distribution are within two standard deviations of the mean. The two points that are two standard deviations from the mean are and so that the interval will contain approximately 95 percent of the possible values of We know that and, hence are unknown, but we may arbitrarily place the sampling distribution of on the -axis. x x x. m mx,

A partir de nuestro conocimiento de las distribuciones normales, en general, sabemos aún más sobre la distribución de en este caso. Sabemos, por ejemplo, que independientemente de dónde se encuentre la distribución de, aproximadamente el 95 % de los posibles valores de que la constituyen se encuentran dentro de dos desviaciones estándar de la media. Los dos puntos que están a dos desviaciones estándar de la media son y, por lo tanto, el intervalo contendrá aproximadamente el 95 % de los posibles valores de. Sabemos que y, por lo tanto, son desconocidos, pero podemos colocar arbitrariamente la distribución muestral de en el eje x x x. m mx.



Since we do not know the value of not a great deal is accomplished by the expression We do, however, have a point estimate of which is Would it be useful to construct an interval about this point estimate of The answer is yes. Suppose we constructed intervals about every possible value of computed from all possible samples of size n from the population of interest. We would have a large number of intervals of the form with widths all equal to the width of the interval about the unknown

Dado que desconocemos el valor de , no se logra mucho con la expresión. Sin embargo, tenemos una estimación puntual de , que es: ¿Sería útil construir un intervalo en torno a esta estimación puntual de ? La respuesta es sí. Supongamos que construimos intervalos en torno a cada valor posible de , calculado a partir de todas las muestras posibles de tamaño n de la población de interés. Tendríamos una gran cantidad de intervalos de la forma con anchos todos iguales al ancho del intervalo en torno a la incógnita.

Approximately 95 percent of these intervals would have centers falling within the interval about Each of the intervals whose centers fall within of would contain These concepts are illustrated in Figure 6.2.1, in which we see that , and all fall within the interval about and, consequently, the intervals about these sample means include the value of The sample means and do not fall within the interval about and the intervals about them do not include

Aproximadamente el 95 por ciento de estos intervalos tendrían centros que caen dentro del intervalo aproximadamente Cada uno de los intervalos cuyos centros caen dentro de contendría Estos conceptos se ilustran en la Figura 6.2.1, en la que vemos que , y todos caen dentro del intervalo aproximadamente y, en consecuencia, los intervalos alrededor de estas medias muestrales incluyen el valor de Las medias muestrales y no caen dentro del intervalo aproximadamente y los intervalos alrededor de ellas no incluyen

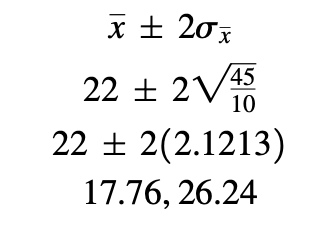
EXAMPLE 6.2.1

Suppose a researcher, interested in obtaining an estimate of the average level of some enzyme in a certain human population, takes a sample of 10 individuals, determines the level of the enzyme in each, and computes a sample mean of Suppose further it is known that the variable of interest is approximately normally distributed with a variance of 45. We wish to estimate

Supongamos que un investigador, interesado en obtener una estimación del nivel promedio de alguna enzima en una determinada población humana, toma una muestra de 10 individuos, determina el nivel de la enzima en cada uno y calcula una media muestral de Supongamos además que se sabe que la variable de interés se distribuye de manera aproximadamente normal con una varianza de 45. Deseamos estimar

**Solution**: An approximate 95 percent confidence interval for is given by

Un intervalo de confianza aproximado del 95 por ciento viene dado por



**Interval Estimate Components**

Let us examine the composition of the interval estimate constructed in Example 6.2.1. It contains in its center the point estimate of The 2 we recognize as a value from the standard normal distribution that tells us within how many standard errors lie approximately 95 percent of the possible values of This value of z is referred to as the reliability coefficient. The last component, is the standard error, or standard deviation of the sampling distribution of In general, then, an interval estimate may be expressed as follows:

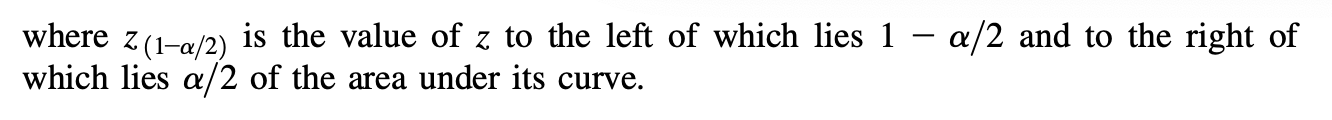
Examinemos la composición de la estimación de intervalo construida en el Ejemplo 6.2.1. Contiene en su centro la estimación puntual de 2. Reconocemos que 2 es un valor de la distribución normal estándar que nos indica dentro de cuántos errores estándar se encuentra aproximadamente el 95% de los valores posibles de . Este valor de z se denomina coeficiente de fiabilidad. El último componente es el error estándar, o desviación estándar, de la distribución muestral de . En general, una estimación de intervalo puede expresarse de la siguiente manera:

estimator +/- (reliability coefficient) x (standard error) (6.2.1)

In particular, when sampling is from a normal distribution with known variance, an interval estimate for may be expressed as



(6.2.2)



where is the value of z to the left of which lies and to the right of which lies of the area under its curve.

Interpreting Confidence Intervals

How do we interpret the interval given by Expression 6.2.2? In the present example, where the reliability coefficient is equal to 2, we say that in repeated sampling approximately 95 percent of the intervals constructed by Expression 6.2.2 will include the population mean. This interpretation is based on the probability of occurrence of different values of We may generalize this interpretation if we designate the total area under the curve of that is outside the interval as and the area within the interval as and give the following probabilistic interpretation of Expression 6.2.2.

¿Cómo interpretamos el intervalo dado por la Expresión 6.2.2? En el presente ejemplo, donde el coeficiente de fiabilidad es igual a 2, decimos que, en el muestreo repetido, aproximadamente el 95 % de los intervalos construidos por la Expresión 6.2.2 incluirán la media poblacional. Esta interpretación se basa en la probabilidad de ocurrencia de diferentes valores de . Podemos generalizar esta interpretación si designamos el área total bajo la curva de que está fuera del intervalo como y el área dentro del intervalo como , y damos la siguiente interpretación probabilística de la Expresión 6.2.2.

**Probabilistic Interpretation**

In repeated sampling, from a normally distributed population with a known standard deviation, percent of all intervals of the form will in the long run include the population mean

En un muestreo repetido, a partir de una población distribuida normalmente con una desviación estándar conocida, el porcentaje de todos los intervalos de la forma incluirá, a largo plazo, la media de la población.

The quantity in this case .95, is called the confidence coefficient (or confidence level), and the interval is called a confidence interval for When the interval is called the 95 percent confidence interval for In the present example we say that we are 95 percent confident that the population mean is between 17.76 and 26.24. This is called the practical interpretation of Expression 6.2.2. In general, it may be expressed as follows.

La cantidad, en este caso 0,95, se denomina coeficiente de confianza (o nivel de confianza), y el intervalo se denomina intervalo de confianza para [número]. Cuando el intervalo se denomina intervalo de confianza del 95 % para [número]. En el presente ejemplo, decimos que tenemos un 95 % de confianza en que la media poblacional está entre 17,76 y 26,24. Esto se denomina interpretación práctica de la expresión 6.2.2. En general, se puede expresar de la siguiente manera.

**Practical Interpretation**

When sampling is from a normally distributed population with known standard deviation, we are percent confident that the single computed interval, , contains the population mean

In the example given here we might prefer, rather than 2, the more exact value of z, 1.96, corresponding to a confidence coefficient of .95. Researchers may use any confidence coefficient they wish; the most frequently used values are .90, .95, and .99, which have associated reliability factors, respectively, of 1.645, 1.96, and 2.58.

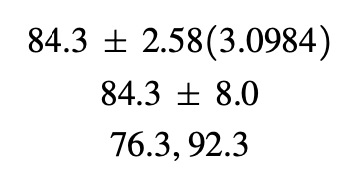
**Precision** The quantity obtained by multiplying the reliability factor by the standard error of the mean is called the precision of the estimate. This quantity is also called the margin of error.

**EXAMPLE 6.2.2**

A physical therapist wished to estimate, with 99 percent confidence, the mean maximal strength of a particular muscle in a certain group of individuals. He is willing to assume that strength scores are approximately normally distributed with a variance of 144. A sample of 15 subjects who participated in the experiment yielded a mean of 84.3.

Un fisioterapeuta deseaba estimar, con un 99 por ciento de confianza, la fuerza máxima media de un músculo en particular en un determinado grupo de individuos. Está dispuesto a suponer que las puntuaciones de fuerza se distribuyen aproximadamente normalmente con una varianza de 144. Una muestra de 15 sujetos que participaron en el experimento arrojó una media de 84,3.

**Solution**: The z value corresponding to a confidence coefficient of .99 is found in Appendix Table D to be 2.58. This is our reliability coefficient. The standard error is Our 99 percent confidence interval for then, is



We say we are 99 percent confident that the population mean is between 76.3 and 92.3 since, in repeated sampling, 99 percent of all intervals that could be constructed in the manner just described would include the population mean. ■

Decimos que estamos 99 por ciento seguros de que la media de la población está entre 76,3 y 92,3 ya que, en un muestreo repetido, el 99 por ciento de todos los intervalos que podrían construirse de la manera recién descrita incluirían la media de la población.

Situations in which the variable of interest is approximately normally distributed with a known variance are so rare as to be almost nonexistent. The purpose of the preceding examples, which assumed that these ideal conditions existed, was to establish the theoretical background for constructing confidence intervals for population means. In most practical situations either the variables are not approximately normally distributed or the population variances are not known or both. Example 6.2.3 and Section 6.3 explain the procedures that are available for use in the less than ideal, but more common, situations.

Las situaciones en las que la variable de interés se distribuye aproximadamente de forma normal con una varianza conocida son tan poco frecuentes que son prácticamente inexistentes. El propósito de los ejemplos anteriores, que asumían la existencia de estas condiciones ideales, fue establecer las bases teóricas para la construcción de intervalos de confianza para las medias poblacionales. En la mayoría de las situaciones prácticas, las variables no se distribuyen aproximadamente de forma normal, se desconocen las varianzas poblacionales, o ambas. El Ejemplo 6.2.3 y la Sección 6.3 explican los procedimientos disponibles para las situaciones menos ideales, pero más comunes.

**Sampling from Nonnormal Populations**

As noted, it will not always be possible or prudent to assume that the population of interest is normally distributed. Thanks to the central limit theorem, this will not deter us if we are able to select a large enough sample. We have learned that for large samples, the sampling distribution of is approximately normally distributed regardless of how the parent population is distributed.

Como se indicó, no siempre será posible ni prudente asumir que la población de interés se distribuye normalmente. Gracias al teorema del límite central, esto no nos disuadirá si logramos seleccionar una muestra suficientemente grande. Hemos aprendido que, para muestras grandes, la distribución muestral de se distribuye aproximadamente de forma normal, independientemente de cómo se distribuya la población original.

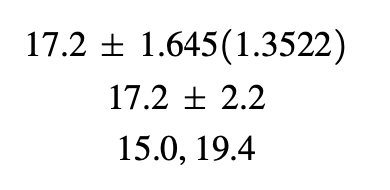
**EXAMPLE 6.2.3**

Punctuality of patients in keeping appointments is of interest to a research team. In a study of patient flow through the offices of general practitioners, it was found that a sample of 35 patients were 17.2 minutes late for appointments, on the average. Previous research had shown the standard deviation to be about 8 minutes. The population distribution was felt to be nonnormal. What is the 90 percent confidence interval for mu, the true mean amount of time late for appointments?

La puntualidad de los pacientes en la asistencia a las citas es de interés para un equipo de investigación. En un estudio sobre el flujo de pacientes en los consultorios de los médicos generales, se encontró que una muestra de 35 pacientes llegaba 17,2 minutos tarde a sus citas, en promedio. Investigaciones anteriores habían demostrado que la desviación estándar era de unos 8 minutos. Se consideró que la distribución de la población no era normal. ¿Cuál es el intervalo de confianza del 90 por ciento para mu, la verdadera cantidad media de retraso en las citas?

Solution: Since the sample size is fairly large (greater than 30), and since the population standard deviation is known, we draw on the central limit theorem and assume the sampling distribution of to be approximately normally distributed. From Appendix Table D we find the reliability coefficient corresponding to a confidence coefficient of .90 to be about 1.645, if we interpolate. The standard error is so that our 90 percent confidence interval for is

Solución: Dado que el tamaño de la muestra es bastante grande (mayor de 30) y se conoce la desviación estándar de la población, nos basamos en el teorema del límite central y asumimos que la distribución muestral de se distribuye aproximadamente de forma normal. De la Tabla D del Apéndice, encontramos que el coeficiente de fiabilidad correspondiente a un coeficiente de confianza de 0,90 es aproximadamente 1,645, si interpolamos. El error estándar es tal que nuestro intervalo de confianza del 90% para es



■

Frequently, when the sample is large enough for the application of the central limit theorem, the population variance is unknown. In that case we use the sample variance as a replacement for the unknown population variance in the formula for constructing a confidence interval for the population mean.

Con frecuencia, cuando la muestra es lo suficientemente grande como para aplicar el teorema del límite central, se desconoce la varianza poblacional. En ese caso, utilizamos la varianza muestral como sustituto de la varianza poblacional desconocida en la fórmula para construir un intervalo de confianza para la media poblacional.

**Computer Analysis**

When confidence intervals are desired, a great deal of time can be saved if one uses a computer, which can be programmed to construct intervals from raw data.

**EXAMPLE 6.2.4**

The following are the activity values (micromoles per minute per gram of tissue) of a certain enzyme measured in normal gastric tissue of 35 patients with gastric carcinoma.

Los siguientes son los valores de actividad (micromoles por minuto por gramo de tejido) de cierta enzima medidos en tejido gástrico normal de 35 pacientes con carcinoma gástrico.

.360, 1.189, .614, .788, .273, 2.464, .571, 1.827, .537, .374, .449, .262, .448, .971, .372, .898, .411, .348, 1.925, .550, .622, .610, .319, .406, .413, .767, .385, .674, .521, .603, .533, .662, 1.177, .307, 1.499

We wish to use the MINITAB computer software package to construct a 95 percent confidence interval for the population mean. Suppose we know that the population variance is .36. It is not necessary to assume that the sampled population of values is normally distributed since the sample size is sufficiently large for application of the central limit theorem.

**Solution**: We enter the data into Column 1 and proceed as shown in Figure 6.2.2. These instructions tell the computer that the reliability factor is z, that a 95 percent confidence interval is desired, that the population standard deviation is .6, and that the data are in Column 1. The output tells us that the sample mean is .718, the sample standard deviation is .511, and the standard error of the mean, is

We are 95 percent confident that the population mean is somewhere between .519 and .917. Confidence intervals may be obtained through the use of many other software packages. Users of SAS®, for example, may wish to use the output from PROC MEANS or PROC UNIVARIATE to construct confidence intervals.

s>1n .6>135 = .101.

15.0, 19.4

17.2 ; 2.2

17.2 ; 1.64511.35222

m

sx = 8> 135 = 1.3522,

x

■

**Alternative Estimates of Central Tendency**

As noted previously, the mean is sensitive to extreme values—those values that deviate appreciably from most of the measurements in a data set. They are sometimes referred to as outliers. We also noted earlier that the median, because it is not so sensitive to extreme measurements, is sometimes preferred over the mean as a measure of central tendency when outliers are present. For the same reason, we may prefer to use the sample median as an estimator of the population median when we wish to make an inference about the central tendency of a population. Not only may we use the sample median as a point estimate of the population median, we also may construct a confidence interval for the population median. The formula is not given here but may be found in the book by Rice (1).

Como se mencionó anteriormente, la media es sensible a los valores extremos, es decir, a aquellos que se desvían considerablemente de la mayoría de las mediciones en un conjunto de datos. A veces se les denomina valores atípicos. También mencionamos anteriormente que la mediana, debido a su menor sensibilidad a las mediciones extremas, a veces se prefiere a la media como medida de tendencia central cuando existen valores atípicos. Por la misma razón, podemos preferir usar la mediana muestral como estimador de la mediana poblacional cuando deseamos inferir la tendencia central de una población. No solo podemos usar la mediana muestral como estimación puntual de la mediana poblacional, sino que también podemos construir un intervalo de confianza para esta. La fórmula no se proporciona aquí, pero puede encontrarse en el libro de Rice (1).

**Trimmed Mean** Estimators that are insensitive to outliers are called robust estimators. Another robust measure and estimator of central tendency is the trimmed mean. For a set of sample data containing n measurements we calculate the percent trimmed mean as follows:

Los estimadores insensibles a valores atípicos se denominan estimadores robustos. Otra medida robusta y estimador de tendencia central es la media recortada. Para un conjunto de datos muestrales con n mediciones, calculamos el porcentaje de la media recortada de la siguiente manera:

1. Order the measurements.

2. Discard the smallest percent and the largest percent of the measurements. The recommended value of is something between .1 and .2.

3. Compute the arithmetic mean of the remaining measurements. Note that the median may be regarded as a 50 percent trimmed mean.

EXERCISES

For each of the following exercises construct 90, 95, and 99 percent confidence intervals for the

population mean, and state the practical and probabilistic interpretations of each. Indicate which

interpretation you think would be more appropriate to use when discussing confidence intervals with

a

100a 100a

100a

EXERCISES 171

FIGURE 6.2.2 MINITAB procedure for constructing 95 percent confidence interval for a

population mean, Example 6.2.4.

Dialog box: Session command:

Stat ➤ Basic Statistics ➤ 1-Sample z MTB > ZINTERVAL 95 .6 C1

Type C1 in Samples in Columns.

Type .6 in Standard deviation. Click OK.

Output:

One-Sample Z: C1

The assumed standard deviaion " 0.600

Variable N Mean StDev SE Mean 95.0 % C.I.

MicMoles 35 0.718 0.511 0.101 ( 0.519, 0.917)

172 CHAPTER 6 ESTIMATION

someone who has not had a course in statistics, and state the reason for your choice. Explain why

the three intervals that you construct are not of equal width. Indicate which of the three intervals

you would prefer to use as an estimate of the population mean, and state the reason for your choice.

6.2.1 We wish to estimate the average number of heartbeats per minute for a certain population. The

average number of heartbeats per minute for a sample of 49 subjects was found to be 90. Assume

that these 49 patients constitute a random sample, and that the population is normally distributed

with a standard deviation of 10.

6.2.2 We wish to estimate the mean serum indirect bilirubin level of 4-day-old infants. The mean for a

sample of 16 infants was found to be 5.98 mg!100 cc. Assume that bilirubin levels in 4-day-old

infants are approximately normally distributed with a standard deviation of 3.5 mg!100 cc.

6.2.3 In a length of hospitalization study conducted by several cooperating hospitals, a random sample

of 64 peptic ulcer patients was drawn from a list of all peptic ulcer patients ever admitted to the

participating hospitals and the length of hospitalization per admission was determined for each.

The mean length of hospitalization was found to be 8.25 days. The population standard deviation

is known to be 3 days.

6.2.4 A sample of 100 apparently normal adult males, 25 years old, had a mean systolic blood pressure

of 125. It is believed that the population standard deviation is 15.

6.2.5 Some studies of Alzheimer’s disease (AD) have shown an increase in production in patients

with the disease. In one such study the following values were obtained from 16 neocorti-

cal biopsy samples from AD patients.

1009 1280 1180 1255 1547 2352 1956 1080

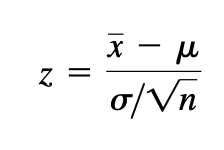
1776 1767 1680 2050 1452 2857 3100 1621

Assume that the population of such values is normally distributed with a standard deviation of 350.

6.3 THE t DISTRIBUTION

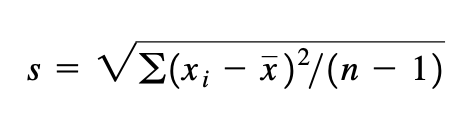
In Section 6.2, a procedure was outlined for constructing a confidence interval for a population mean. The procedure requires knowledge of the variance of the population from which the sample is drawn. It may seem somewhat strange that one can have knowledge of the population variance and not know the value of the population mean. Indeed, it is the usual case, in situations such as have been presented, that the population variance, as well as the population mean, is unknown. This condition presents a problem with respect to constructing confidence intervals. Although, for example, the statistic

En la Sección 6.2, se describió un procedimiento para construir un intervalo de confianza para una media poblacional. El procedimiento requiere conocer la varianza de la población de la que se extrae la muestra. Puede parecer extraño que se conozca la varianza poblacional y se desconozca el valor de la media poblacional. De hecho, es habitual, en situaciones como las presentadas, que se desconozcan tanto la varianza poblacional como la media poblacional. Esta condición presenta un problema con respecto a la construcción de intervalos de confianza. Aunque, por ejemplo, el estadístico



is normally distributed when the population is normally distributed and is at least approximately normally distributed when n is large, regardless of the functional form of the population, we cannot make use of this fact because is unknown. However, all is not lost, and the most logical solution to the problem is the one followed. We use the sample standard deviation

Se distribuye normalmente cuando la población también lo hace, y se distribuye al menos aproximadamente de forma normal cuando n es grande. Independientemente de la forma funcional de la población, no podemos aprovechar este hecho porque se desconoce. Sin embargo, no todo está perdido, y la solución más lógica al problema es la siguiente: utilizamos la desviación estándar muestral.



to replace When the sample size is large, say, greater than 30, our faith in s as an approximation of is usually substantial, and we may be appropriately justified in using normal distribution theory to construct a confidence interval for the population mean. In that event, we proceed as instructed in Section 6.2.

It is when we have small samples that it becomes mandatory for us to find an alter- native procedure for constructing confidence intervals.

As a result of the work of Gosset (2), writing under the pseudonym of “Student,” an alternative, known as Student’s t distribution, usually shortened to t distribution, is available to us.

The quantity



(6.3.1)

follows this distribution.

**Properties of the t Distribution**

The t distribution has the following

properties.

1. It has a mean of 0.

2. It is symmetrical about the mean.

3. In general, it has a variance greater than 1, but the variance approaches 1 as the sample size becomes large. For the variance of the t distribution is where df is the degrees of freedom. Alternatively, since here for we may write the variance of the t distribution as

4. The variable t ranges from to

5. The t distribution is really a family of distributions, since there is a different distribution for each sample value of the divisor used in computing We recall that is referred to as degrees of freedom. Figure 6.3.1 shows t distributions corresponding to several degrees-of-freedom values.

n - 1

s2 n - 1, .

- q + q.

1n - 12>1n - 32.

df = n - 1 n 7 3,

df>1df - 22,

df 7 2,

t = x - m

s>1n

s

s.

6.3 THE t DISTRIBUTION 173

Degrees of freedom = 30

Degrees of freedom = 5

Degrees of freedom = 2

t

FIGURE 6.3.1 The t distribution for different degrees-of-freedom values.

174 CHAPTER 6 ESTIMATION

6. Compared to the normal distribution, the t distribution is less peaked in the center and has thicker tails. Figure 6.3.2 compares the t distribution with the normal.

7. The t distribution approaches the normal distribution as approaches infinity.

The t distribution, like the standard normal, has been extensively tabulated. One such table is given as Table E in the Appendix. As we will see, we must take both the confidence coefficient and degrees of freedom into account when using the table of the t distribution.

You may use MINITAB to graph the t distribution (for specified degrees-of-freedom values) and other distributions. After designating the horizontal axis by following directions in the Set Patterned Data box, choose menu path Calc and then Probability Distributions. Finally, click on the distribution desired and follow instructions. Use the Plot dialog box to plot the graph.

**Confidence Intervals Using t**

The general procedure for constructing confidence intervals is not affected by our having to use the t distribution rather than the standard normal distribution. We still make use of the relationship expressed by

reliability coefficient

What is different is the source of the reliability coefficient. It is now obtained from the table of the t distribution rather than from the table of the standard normal distribution. To be more specific, when sampling is from a normal distribution whose standard deviation, is unknown, the percent confidence interval for the population mean, is given by



(6.3.2)

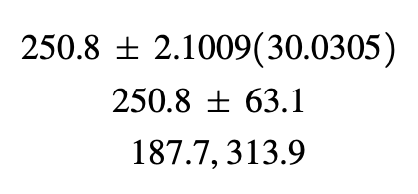
We emphasize that a requirement for the strictly valid use of the t distribution is that the sample must be drawn from a normal distribution. Experience has shown, however, that moderate departures from this requirement can be tolerated. As a consequence, the t distribution is used even when it is known that the parent population deviates somewhat from normality. Most researchers require that an assumption of, at least, a mound-shaped population distribution be tenable.

**EXAMPLE 6.3.1**

Maffulli et al. (A-1) studied the effectiveness of early weightbearing and ankle mobilization therapies following acute repair of a ruptured Achilles tendon. One of the variables they measured following treatment was the isometric gastrocsoleus muscle strength. In 19 subjects, the mean isometric strength for the operated limb (in newtons) was 250.8 with a standard deviation of 130.9. We assume that these 19 patients constitute a random sample from a population of similar subjects. We wish to use these sample data to estimate for the population the mean isometric strength after surgery.

Maffulli et al. (A-1) estudiaron la eficacia de las terapias tempranas con soporte de peso y movilización del tobillo después de la reparación aguda de una rotura del tendón de Aquiles. Una de las variables que midieron después del tratamiento fue la fuerza isométrica del músculo gastrocsoleo. En 19 sujetos, la fuerza isométrica media de la extremidad operada (en newtons) fue de 250,8 con una desviación estándar de 130,9. Suponemos que estos 19 pacientes constituyen una muestra aleatoria de una población de sujetos similares. Deseamos utilizar estos datos de muestra para estimar para la población la fuerza isométrica media después de la cirugía.

**Solution**: We may use the sample mean, 250.8, as a point estimate of the population mean but, because the population standard deviation is unknown, we must assume the population of values to be at least approximately normally distributed before constructing a confidence interval for Let us assume that such an assumption is reasonable and that a 95 percent confidence interval is desired. We have our estimator, and our standard error is We need now to find the reliability coefficient, the value of t associated with a confidence coefficient of .95 and degrees of freedom. Since a 95 percent confidence interval leaves .05 of the area under the curve of t to be equally divided between the two tails, we need the value of t to the right of which lies .025 of the area. We locate in Appendix Table E the column headed This is the value of t to the left of which lies .975 of the area under the curve. The area to the right of this value is equal to the desired .025. We now locate the number 18 in the degrees-of-freedom column. The value at the intersection of the row labeled 18 and the column labeled is the t we seek. This value of t, which is our reliability coefficient, is found to be 2.1009. We now construct our 95 percent confidence interval as follows:



■

This interval may be interpreted from both the probabilistic and practical points of view. We are 95 percent confident that the true population mean, is somewhere between 187.7 and 313.9 because, in repeated sampling, 95 percent of intervals constructed in like manner will include

**Deciding Between z and t**

When we construct a confidence interval for a population mean, we must decide whether to use a value of z or a value of t as the reliability factor. To make an appropriate choice we must consider sample size, whether the sampled population is normally distributed, and whether the population variance is known. Figure 6.3.3 provides a flowchart that one can use to decide quickly whether the reliability factor should be z or t.

**Computer Analysis**

If you wish to have MINITAB construct a confidence interval for a population mean when the t statistic is the appropriate reliability factor, the command is TINTERVAL. In Windows choose 1-Sample t from the Basic Statistics menu.

**EXERCISES**

6.3.1 Use the t distribution to find the reliability factor for a confidence interval based on the following

confidence coefficients and sample sizes:

abcd

Confidence coefficient .95 .99 .90 .95

Sample size 15 24 8 30

6.3.2 In a study of the effects of early Alzheimer’s disease on nondeclarative memory, Reber et al. (A-2)

used the Category Fluency Test to establish baseline persistence and semantic memory and language

abilities. The eight subjects in the sample had Category Fluency Test scores of 11, 10, 6, 3, 11, 10,

9, 11. Assume that the eight subjects constitute a simple random sample from a normally distributed

population of similar subjects with early Alzheimer’s disease.

(a) What is the point estimate of the population mean?

(b) What is the standard deviation of the sample?

(c) What is the estimated standard error of the sample mean?

(d) Construct a 95 percent confidence interval for the population mean category fluency test score.

(e) What is the precision of the estimate?

(f) State the probabilistic interpretation of the confidence interval you constructed.

(g) State the practical interpretation of the confidence interval you constructed.

6.3.3 Pedroletti et al. (A-3) reported the maximal nitric oxide diffusion rate in a sample of 15 asthmatic

schoolchildren and 15 controls as mean standard error of the mean. For asthmatic children, they ;

Population

normally

distributed

Population

variance

known?

Population

variance

known?

Population

variance

known?

Population

normally

distributed?

Yes

Yes

No Yes No

No Yes No

or

Yes

Yes

No

\*

Yes No

No

Sample

size

large?

Sample

size

large?

Population

variance

known?

z

z

z t tz z

Central limit theorem applies

\*

FIGURE 6.3.3 Flowchart for use in deciding between z and t when making inferences

about population means. (\*Use a nonparametric procedure. See Chapter 13.)

reported (nanoliters per second) and for control subjects they reported

For each group, determine the following:

(a) What was the sample standard deviation?

(b) What is the 95 percent confidence interval for the mean maximal nitric oxide diffusion rate

of the population?

(c) What assumptions are necessary for the validity of the confidence interval you constructed?

(d) What are the practical and probabilistic interpretations of the interval you constructed?

(e) Which interpretation would be more appropriate to use when discussing confidence intervals

with someone who has not had a course in statistics? State the reasons for your choice.

(f) If you were to construct a 90 percent confidence interval for the population mean from the

information given here, would the interval be wider or narrower than the 95 percent confidence

interval? Explain your answer without actually constructing the interval.

(g) If you were to construct a 99 percent confidence interval for the population mean from the

information given here, would the interval be wider or narrower than the 95 percent confidence

interval? Explain your answer without actually constructing the interval.

6.3.4 The concern of a study by Beynnon et al. (A-4) were nine subjects with chronic anterior cru-

ciate ligament (ACL) tears. One of the variables of interest was the laxity of the anteroposte-

rior, where higher values indicate more knee instability. The researchers found that among

subjects with ACL-deficient knees, the mean laxity value was 17.4 mm with a standard devi-

ation of 4.3 mm.

(a) What is the estimated standard error of the mean?

(b) Construct the 99 percent confidence interval for the mean of the population from which the

nine subjects may be presumed to be a random sample.

(c) What is the precision of the estimate?

(d) What assumptions are necessary for the validity of the confidence interval you constructed?

6.3.5 A sample of 16 ten-year-old girls had a mean weight of 71.5 and a standard deviation of 12 pounds,

respectively. Assuming normality, find the 90, 95, and 99 percent confidence intervals for

6.3.6 The subjects of a study by Dugoff et al. (A-5) were 10 obstetrics and gynecology interns at the

University of Colorado Health Sciences Center. The researchers wanted to assess competence in

performing clinical breast examinations. One of the baseline measurements was the number of

such examinations performed. The following data give the number of breast examinations per-

formed for this sample of 10 interns.

Intern Number No. of Breast Exams Performed

1 30

2 40

3 8

4 20

5 26

6 35

7 35

8 20

9 25

10 20

m.

0.7 ; .1 nL/s.

3.5 ; 0.4 nL/s

EXERCISES 177

Source: Lorraine Dugoff, Mauritha R.

Everett, Louis Vontver, and Gwyn E.

Barley, “Evaluation of Pelvic and Breast

Examination Skills of Interns in Obstetrics

and Gynecology and Internal Medicine,”

American Journal of Obstetrics and

Gynecology, 189 (2003), 655–658.

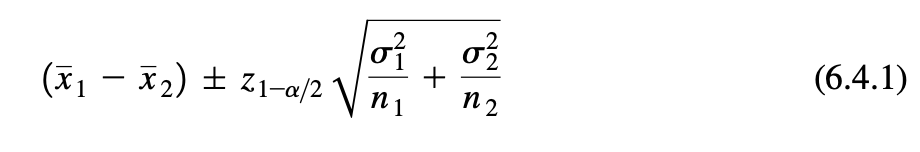
178 CHAPTER 6 ESTIMATION

Construct a 95 percent confidence interval for the mean of the population from which the study

subjects may be presumed to have been drawn.

6.4 CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN TWO POPULATION MEANS

Sometimes there arise cases in which we are interested in estimating the difference between two population means. From each of the populations an independent random sample is drawn and, from the data of each, the sample means and respectively, are computed. We learned in the previous chapter that the estimator yields an unbiased estimate of the difference between the population means. The variance of the estimator is We also know from Chapter 5 that, depending on the conditions, the sampling distribution of may be, at least, approximately normally distributed, so that in many cases we make use of the theory relevant to normal distributions to compute a confidence interval for When the population variances are known, the percent confidence interval for is given by



(6.4.1)

An examination of a confidence interval for the difference between population means provides information that is helpful in deciding whether or not it is likely that the two population means are equal. When the constructed interval does not include zero, we say that the interval provides evidence that the two population means are not equal. When the interval includes zero, we say that the population means may be equal. Let us illustrate for the case where sampling is from normal distributions.

**EXAMPLE 6.4.1**

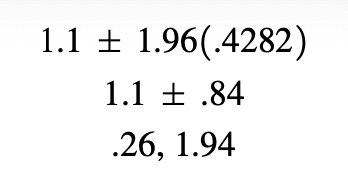
A research team is interested in the difference between serum uric acid levels in patients with and without Down’s syndrome. In a large hospital for the treatment of the mentally retarded, a sample of 12 individuals with Down’s syndrome yielded a mean of In a general hospital a sample of 15 normal individuals of the same age and sex were found to have a mean value of If it is reasonable to assume that the two populations of values are normally distributed with variances equal to 1 and 1.5, find the 95 percent confidence interval for

**Solution**: For a point estimate of we use The reliability coefficient corresponding to .95 is found in Appendix Table D to be 1.96. The standard error is

sx1-x2 = C

x 2 = 3.4.

The 95 percent confidence interval, then, is



We say that we are 95 percent confident that the true difference, is somewhere between .26 and 1.94 because, in repeated sampling, 95 percent of the intervals constructed in this manner would include the difference between the true means. Since the interval does not include zero, we conclude that the two population means are not equal. ■

**Sampling from Nonnormal Populations**

The construction of a confidence interval for the difference between two population means when sampling is from nonnormal populations proceeds in the same manner as in Example 6.4.1 if the sample sizes and are large. Again, this is a result of the central limit theorem. If the population variances are unknown, we use the sample variances to estimate them.

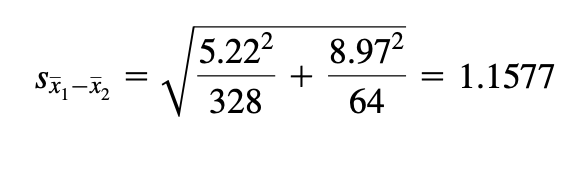
**EXAMPLE 6.4.2**

Despite common knowledge of the adverse effects of doing so, many women continue to smoke while pregnant. Mayhew et al. (A-6) examined the effectiveness of a smoking cessation program for pregnant women. The mean number of cigarettes smoked daily at the close of the program by the 328 women who completed the program was 4.3 with a standard deviation of 5.22. Among 64 women who did not complete the program, the mean number of cigarettes smoked per day at the close of the program was 13 with a standard deviation of 8.97. We wish to construct a 99 percent confidence interval for the difference between the means of the populations from which the samples may be pre-

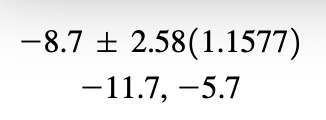
sumed to have been selected.

**Solution**: No information is given regarding the shape of the distribution of cigarettes smoked per day. Since our sample sizes are large, however, the central limit theorem assures us that the sampling distribution of the difference between sample means will be approximately normally distributed even if the distribution of the variable in the populations is not normally distributed. We may use this fact as justification for using the z statistic as the reliability factor

in the construction of our confidence interval. Also, since the population standard deviations are not given, we will use the sample standard deviations to estimate them. The point estimate for the difference between population means is the difference between sample means, In Appendix Table D we find the reliability factor to be 2.58. The estimated standard error is



By Equation 6.4.1, our 99 percent confidence interval for the difference between population means is



We are 99 percent confident that the mean number of cigarettes smoked per day for women who complete the program is between 5.7 and 11.7 lower than the mean for women who do not complete the program. ■

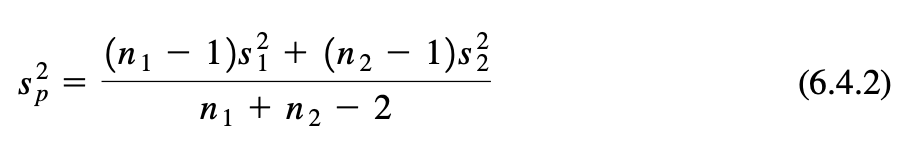
**The t Distribution and the Difference Between Means**

When population variances are unknown, and we wish to estimate the difference between two population means with a confidence interval, we can use the t distribution as a source of the reliability factor if certain assumptions are met. We must know, or be willing to assume, that the two sampled populations are normally distributed. With regard to the population variances, we distinguish between two situations: (1) the situation in which the population variances are equal, and (2) the situation in which they are not equal. Let us consider each situation separately.

**Population Variances Equal**

Variaciones de población iguales

If the assumption of equal population variances is justified, the two sample variances that we compute from our two independent samples may be considered as estimates of the same quantity, the common variance. It seems logical, then, that we should somehow capitalize on this in our analysis. We do just tha and obtain a pooled estimate of the common variance. This pooled estimate is obtained by computing the weighted average of the two sample variances. Each sample variance is weighted by its degrees of freedom. If the sample sizes are equal, this weighted average is the arithmetic mean of the two sample variances. If the two sample sizes are unequal, the weighted average takes advantage of the additional information provided by the larger sample. The pooled estimate is given by the formula

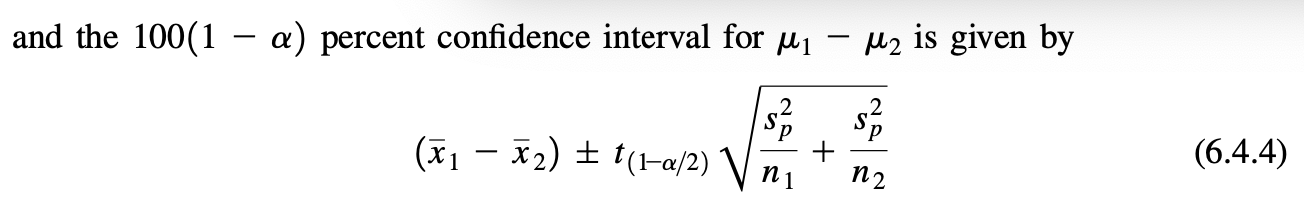


(6.4.2)

The standard error of the estimate, then, is given by



(6.4.3)



and the percent confidence interval for is given by

(6.4.4)

The number of degrees of freedom used in determining the value of t to use in constructing the interval is n1 + n2 -2 the denominator of Equation 6.4.2. We interpret this interval in the usual manner.

Methods that may be used in reaching a decision about the equality of population

variances are discussed in Sections 6.10 and 7.8.

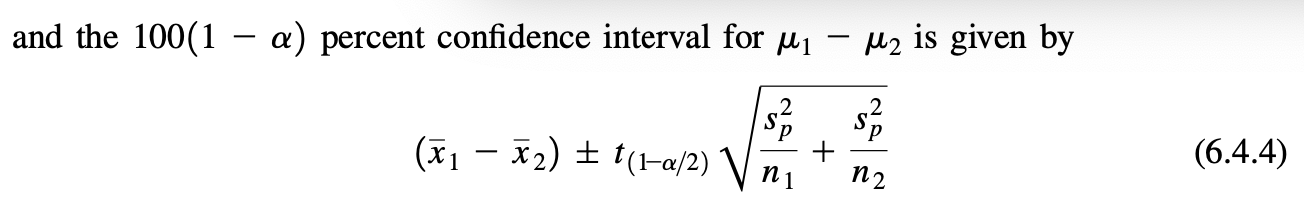
EXAMPLE 6.4.3

The purpose of a study by Granholm et al. (A-7) was to determine the effectiveness of an integrated outpatient dual-diagnosis treatment program for mentally ill subjects. The authors were addressing the problem of substance abuse issues among people with severe mental disorders. A retrospective chart review was carried out on 50 consecutive patient referrals to the Substance Abuse/Mental Illness program at the VA San Diego Healthcare System. One of the outcome variables examined was the number of inpatient treatment days for psychiatric disorder during the year following the end of the program. Among 18 subjects with schizophrenia, the mean number of treatment days was 4.7 with a standard deviation of 9.3. For 10 subjects with bipolar disorder, the mean number of psychiatric disorder treatment days was 8.8 with a standard deviation of 11.5. We wish to construct a 95 percent confidence interval for the difference between the means of the populations represented by these two samples.

El propósito de un estudio de Granholm et al. (A-7) fue determinar la efectividad de un programa de tratamiento ambulatorio integrado de diagnóstico dual para sujetos con enfermedades mentales. Los autores abordaban el problema del abuso de sustancias entre personas con trastornos mentales graves. Se llevó a cabo una revisión retrospectiva de los expedientes de 50 pacientes remitidos consecutivamente al programa de Abuso de Sustancias/Enfermedades Mentales del Sistema de Atención Médica VA de San Diego. Una de las variables de resultado examinadas fue el número de días de tratamiento hospitalario por trastorno psiquiátrico durante el año siguiente a la finalización del programa. Entre 18 sujetos con esquizofrenia, el número medio de días de tratamiento fue de 4,7 con una desviación estándar de 9,3. Para 10 sujetos con trastorno bipolar, el número medio de días de tratamiento del trastorno psiquiátrico fue de 8,8 con una desviación estándar de 11,5. Deseamos construir un intervalo de confianza del 95 por ciento para la diferencia entre las medias de las poblaciones representadas por estas dos muestras.

**Solution**: First we use Equation 6.4.2 to compute the pooled estimate of the common population variance.

When we enter Appendix Table E with degrees of freedom and a desired confidence level of .95, we find that the reliability factor is 2.0555. By Expression 6.4.4 we compute the 95 percent confidence interval for the difference between population means as follows:



We are 95 percent confident that the difference between population means is somewhere between and 4.10. We can say this because we know that if we were to repeat the study many, many times, and compute confidence intervals in the same way, about 95 percent of the intervals would include the difference between the population means.

Since the interval includes zero, we conclude that the population means may be equal. ■

**Population Variances Not Equal**

When one is unable to conclude that the variances of two populations of interest are equal, even though the two populations may be assumed to be normally distributed, it is not proper to use the t distribution as just outlined in constructing confidence intervals.

A solution to the problem of unequal variances was proposed by Behrens (3) and later was verified and generalized by Fisher (4, 5). Solutions have also been proposed by Neyman (6), Scheffé (7, 8), and Welch (9, 10). The problem is discussed in detail by Cochran (11).

10

18 + 10 - 2 = 26

s 2

p = 118 - 1219.32

2 + 110 - 12111.522

18 + 10 - 2 = 102.33

6.4 CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN TWO POPULATION MEANS 181

182 CHAPTER 6 ESTIMATION

The problem revolves around the fact that the quantity

does not follow a t distribution with degrees of freedom when the population variances are not equal. Consequently, the t distribution cannot be used in the usual way to obtain the reliability factor for the confidence interval for the difference between the means of two populations that have unequal variances. The solution proposed by Cochran consists of computing the reliability factor, by the following formula:

(6.4.5)

where for degrees of freedom, and for degrees of freedom. An approximate percent confidence interval for is given by

(6.4.6)

Adjustments to the reliability coefficient may also be made by reducing the number of degrees of freedom instead of modifying t in the manner just demonstrated. Many computer programs calculate an adjusted reliability coefficient in this way.

**EXAMPLE 6.4.4**

Let us reexamine the data presented in Example 6.4.3 from the study by Granholm et al. (A-7). Recall that among the 18 subjects with schizophrenia, the mean number of treatment days was 4.7 with a standard deviation of 9.3. In the bipolar disorder treatment group of 10 subjects, the mean number of psychiatric disorder treatment days was 8.8 with a standard deviation of 11.5. We assume that the two populations of number of psychiatric disorder days are approximately normally distributed. Now let us assume, however, that the two population variances are not equal. We wish to construct a 95 percent confidence interval for the difference between the means of the two populations represented by the samples.

Solution: We will use as found in Equation 6.4.5 for the reliability factor. Reference to Appendix Table E shows that with 17 degrees of freedom and

Similarly, with 9 degrees of freedom

and We now compute

t ¿ = 19.32

>18212.10982 + 111.52

>10212.26222

19.32

>182 + 111.52

>102 = 2.2216

1 - .05>2 = .975, t 2 = 2.2622.

1 - .05>2 = .975, t1 = 2.1098.

t ¿

1x 1 - x 22 ; t ¿

11-a>22 A

s 2

1

n1

+

s 2

2

n2

m1 - m2

t1-a>2 n2 - 1 10011 - a2

t w2 = s t1 = t1-a>2 n1 - 1 2 = 2

2>n2 w , 1 = s 2

1>n1,

t ¿

1-a>2 = w1t1 + w2t 2

w1 + w2

t¿

1-a>2,

n1 + n2 - 2

1x 1 - x 22 - 1m1 - m22

B

s 2

p

n 1

+

s 2

p

n2

By Expression 6.4.6 we now construct the 95 percent confidence interval for

the difference between the two population means.

Since the interval does include zero, we conclude that the two population

means may be equal. ■

When constructing a confidence interval for the difference between two population

means one may use Figure 6.4.1 to decide quickly whether the reliability factor should

be or

**EXERCISES**

For each of the following exercises construct 90, 95, and 99 percent confidence intervals for the

difference between population means. Where appropriate, state the assumptions that make your

method valid. State the practical and probabilistic interpretations of each interval that you con-

struct. Consider the variables under consideration in each exercise, and state what use you think

researchers might make of your results.

6.4.1 Iannelo et al. (A-8) performed a study that examined free fatty acid concentrations in 18 lean sub-

jects and 11 obese subjects. The lean subjects had a mean level of 299 Eq/L with a standard m

z, t, t ¿.

-13.5, 5.3

14.7 - 8.82 ; 2.221614.2461752

14.7 - 8.82 ; 2.2216 A

9.32

18

+

11.52

10

EXERCISES 183

Population

normally

distributed?

Yes

Yes

Yes = ?

Yes

Sample

sizes

large?

Population

variances

known?

z

Population

variances

known?

No Yes No

No Yes = ? No Yes = ? No Yes = ? No

z

z z

t t' z z t t'

No Yes

Yes

Yes = ?

No

Sample

sizes

large?

Population

variances

known?

z

Population

variances

known?

No Yes No

No Yes = ? No Yes = ? No Yes = ? No

z z t'

Central limit theorem applies or or

No

\* \*\* \*

FIGURE 6.4.1 Flowchart for use in deciding whether the reliability factor should be z, t, or

when making inferences about the difference between two population means. (\*Use a

nonparametric procedure. See Chapter 13.)

t¿

error of the mean of 30, while the obese subjects had a mean of 744 Eq/L with a standard error

of the mean of 62.

6.4.2 Chan et al. (A-9) developed a questionnaire to assess knowledge of prostate cancer. There was a

total of 36 questions to which respondents could answer “agree,” “disagree,” or “don’t know.”

Scores could range from 0 to 36. The mean scores for Caucasian study participants was 20.6 with

a standard deviation of 5.8, while the mean scores for African-American men was 17.4 with a

standard deviation of 5.8. The number of Caucasian study participants was 185, and the number

of African-Americans was 86.

6.4.3 The objectives of a study by van Vollenhoven et al. (A-10) were to examine the effectiveness of

etanercept alone and etanercept in combination with methotrexate in the treatment of rheumatoid

arthritis. The researchers conducted a retrospective study using data from the STURE database,

which collects efficacy and safety data for all patients starting biological treatments at the major

hospitals in Stockholm, Sweden. The researchers identified 40 subjects who were prescribed etan-

ercept only and 57 subjects who were given etanercept with methotrexate. Using a 100-mm visual

analogue scale (the higher the value, the greater the pain), researchers found that after 3 months

of treatment, the mean pain score was 36.4 with a standard error of the mean of 5.5 for subjects

taking etanercept only. In the sample receiving etanercept plus methotrexate, the mean score was

30.5 with a standard error of the mean of 4.6.

6.4.4 The purpose of a study by Nozawa et al. (A-11) was to determine the effectiveness of segmental

wire fixation in athletes with spondylolysis. Between 1993 and 2000, 20 athletes (6 women and

14 men) with lumbar spondylolysis were treated surgically with the technique. The following table

gives the Japanese Orthopaedic Association (JOA) evaluation score for lower back pain syndrome

for men and women prior to the surgery. The lower score indicates less pain.

Gender JOA scores

Female 14, 13, 24, 21, 20, 21

Male 21, 26, 24, 24, 22, 23, 18, 24, 13, 22, 25, 23, 21, 25

Source: Satoshi Nozawa, Katsuji Shimizu, Kei Miyamoto, and Mizuo

Tanaka, “Repair of Pars Interarticularis Defect by Segmental Wire Fixa-

tion in Young Athletes with Spondylolysis,” American Journal of Sports

Medicine, 31 (2003), 359–364.

6.4.5 Krantz et al. (A-12) investigated dose-related effects of methadone in subjects with torsade de

pointes, a polymorphic ventricular tachycardia. In the study of 17 subjects, nine were being

treated with methadone for opiate dependency and eight for chronic pain. The mean daily dose

of methadone in the opiate dependency group was 541 mg/day with a standard deviation of

156, while the chronic pain group received a mean dose of 269 mg/day with a standard devi-

ation of 316.

6.4.6 Transverse diameter measurements on the hearts of adult males and females gave the following

results:

Group Sample Size (cm) s (cm)

Males 12 13.21 1.05

Females 9 11.00 1.01

Assume normally distributed populations with equal variances.

x

m

184 CHAPTER 6 ESTIMATION

6.5 CONFIDENCE INTERVAL FOR A POPULATION PROPORTION 185

6.4.7 Twenty-four experimental animals with vitamin D deficiency were divided equally into two groups.

Group 1 received treatment consisting of a diet that provided vitamin D. The second group was

not treated. At the end of the experimental period, serum calcium determinations were made with

the following results:

Treated group:

Untreated group:

Assume normally distributed populations with equal variances.

6.4.8 Two groups of children were given visual acuity tests. Group 1 was composed of 11 children

who receive their health care from private physicians. The mean score for this group was 26

with a standard deviation of 5. Group 2 was composed of 14 children who receive their health

care from the health department, and had an average score of 21 with a standard deviation of

6. Assume normally distributed populations with equal variances.

6.4.9 The average length of stay of a sample of 20 patients discharged from a general hospital was

7 days with a standard deviation of 2 days. A sample of 24 patients discharged from a chronic

disease hospital had an average length of stay of 36 days with a standard deviation of 10 days.

Assume normally distributed populations with unequal variances.

6.4.10 In a study of factors thought to be responsible for the adverse effects of smoking on human repro-

duction, cadmium level determinations (nanograms per gram) were made on placenta tissue of a

sample of 14 mothers who were smokers and an independent random sample of 18 nonsmoking

mothers. The results were as follows:

Nonsmokers: 10.0, 8.4, 12.8, 25.0, 11.8, 9.8, 12.5, 15.4, 23.5,

9.4, 25.1, 19.5, 25.5, 9.8, 7.5, 11.8, 12.2, 15.0

Smokers: 30.0, 30.1, 15.0, 24.1, 30.5, 17.8, 16.8, 14.8,

13.4, 28.5, 17.5, 14.4, 12.5, 20.4

Does it appear likely that the mean cadmium level is higher among smokers than nonsmokers?

Why do you reach this conclusion?

6.5 CONFIDENCE INTERVAL FOR A POPULATION PROPORTION

Many questions of interest to the health worker relate to population proportions. What proportion of patients who receive a particular type of treatment recover? What proportion of some population has a certain disease? What proportion of a population is immune to a certain disease?

To estimate a population proportion we proceed in the same manner as when estimating a population mean. A sample is drawn from the population of interest, and the sample proportion, is computed. This sample proportion is used as the point estimator of the population proportion. A confidence interval is obtained by the general formula

In the previous chapter we saw that when both np and are greater than

5, we may consider the sampling distribution of to be quite close to the normal pN

n11 - p2

estimator ; 1reliability coefficient2 \* 1standard error of the estimator2

pN,

x = 7.8 mg>100 ml, s = 2.0

x = 11.1 mg>100 ml, s = 1.5

distribution. When this condition is met, our reliability coefficient is some value of z from the standard normal distribution. The standard error, we have seen, is equal to Since p, the parameter we are trying to estimate, is unknown, we must use as an estimate. Thus, we estimate by , and our percent confidence interval for p is given by

(6.5.1)

We give this interval both the probabilistic and practical interpretations.

**EXAMPLE 6.5.1**

The Pew Internet and American Life Project (A-13) reported in 2003 that 18 percent of Internet users have used it to search for information regarding experimental treatments or medicines. The sample consisted of 1220 adult Internet users, and information was collected from telephone interviews. We wish to construct a 95 percent confidence interval for the proportion of Internet users in the sampled population who have searched for information on experimental treatments or medicines.

**Solution**: We shall assume that the 1220 subjects were sampled in random fashion. The best point estimate of the population proportion is The size of the sample and our estimate of p are of sufficient magnitude to justify use of the standard normal distribution in constructing a confidence interval. The reliability coefficient corresponding to a confidence level of .95 is 1.96, and our estimate of the standard error is

The 95 percent confidence interval for p,

based on these data, is

We are 95 percent confident that the population proportion p is between .158

and .202 because, in repeated sampling, about 95 percent of the intervals con-

structed in the manner of the present single interval would include the true p.

On the basis of these results we would expect, with 95 percent confidence, to

find somewhere between 15.8 percent and 20.2 percent of adult Internet users

to have used it for information on medicine or experimental treatments. ■

**EXERCISES**

For each of the following exercises state the practical and probabilistic interpretations of the inter-

val that you construct. Identify each component of the interval: point estimate, reliability coeffi-

cient, and standard error. Explain why the reliability coefficients are not the same for all exercises.

6.5.1 Luna et al. (A-14) studied patients who were mechanically ventilated in the intensive care unit

of six hospitals in Buenos Aires, Argentina. The researchers found that of 472 mechanically

.158, .202

.18 ; .022

.18 ; 1.961.01102

21.1821.822>1220 = .0110.

spN 2pN11 - pN2>n =

pN = .18.

pN ; z 1-a>22pN11 - pN2>n

10011 - a2

pN spN 2pN11 - pN2>n

spN = 1p11 - p2>n.

186 CHAPTER 6 ESTIMATION

6.6 CONFIDENCE INTERVAL FOR THE DIFFERENCE BETWEEN TWO POPULATION PROPORTIONS 187

ventilated patients, 63 had clinical evidence of ventilator-associated pneumonia (VAP). Construct

a 95 percent confidence interval for the proportion of all mechanically ventilated patients at these

hospitals who may be expected to develop VAP.

6.5.2 Q waves on the electrocardiogram, according to Schinkel et al. (A-15), are often considered to be

reflective of irreversibly scarred myocardium. These researchers assert, however, that there are

some indications that residual viable tissue may be present in Q-wave-infarcted regions. Their study

of 150 patients with chronic electrocardiographic Q-wave infarction found 202 dysfunctional

Q-wave regions. With dobutamine stress echocardiography (DSE), they noted that 118 of these

202 regions were viable with information from the DSE testing. Construct a 90 percent confidence

interval for the proportion of viable regions that one might expect to find a population of dysfunc-

tional Q-wave regions.

6.5.3 In a study by von zur Muhlen et al. (A-16), 136 subjects with syncope or near syncope were stud-

ied. Syncope is the temporary loss of consciousness due to a sudden decline in blood flow to the

brain. Of these subjects, 75 also reported having cardiovascular disease. Construct a 99 percent

confidence interval for the population proportion of subjects with syncope or near syncope who

also have cardiovascular disease.

6.5.4 In a simple random sample of 125 unemployed male high-school dropouts between the ages of 16

and 21, inclusive, 88 stated that they were regular consumers of alcoholic beverages. Construct a

95 percent confidence interval for the population proportion.

6.6 CONFIDENCE INTERVAL FOR

THE DIFFERENCE BETWEEN TWO

POPULATION PROPORTIONS

The magnitude of the difference between two population proportions is often of interest. We may want to compare, for example, men and women, two age groups, two socioeconomic groups, or two diagnostic groups with respect to the proportion possessing some characteristic of interest. An unbiased point estimator of the difference between two population proportions is provided by the difference between sample proportions, As we have seen, when and are large and the population

proportions are not too close to 0 or 1, the central limit theorem applies and normal

distribution theory may be employed to obtain confidence intervals. The standard error

of the estimate usually must be estimated by

because, as a rule, the population proportions are unknown. A percent con-

fidence interval for is given by

(6.6.1)

We may interpret this interval from both the probabilistic and practical points of view.

1pN 1 - pN 22 ; z 1-a>2

A

pN 111 - pN 12

n1

+

p2N 11 - pN22

n2

p1 - p2

10011 - a2

sN pN1-pN2 = A

pN 111 - pN 12

n1

+

p2N 11 - pN22

n2

p1 n1 n2 N - p2N .

EXAMPLE 6.6.1

Connor et al. (A-17) investigated gender differences in proactive and reactive aggression

in a sample of 323 children and adolescents (68 females and 255 males). The subjects

were from unsolicited consecutive referrals to a residential treatment center and a pedi-

atric psychopharmacology clinic serving a tertiary hospital and medical school. In the

sample, 31 of the females and 53 of the males reported sexual abuse. We wish to con-

struct a 99 percent confidence interval for the difference between the proportions of sex-

ual abuse in the two sampled populations.

Solution: The sample proportions for the females and males are, respectively,

and The difference between sample

proportions is The estimated standard

error of the difference between sample proportions is

The reliability factor from Appendix Table D is 2.58, so that our confidence

interval, by Expression 6.6.1, is

We are 99 percent confident that for the sampled populations, the proportion

of cases of reported sexual abuse among females exceeds the proportion of

cases of reported sexual abuse among males by somewhere between .0791

and .4171.

Since the interval does not include zero, we conclude that the two

population proportions are not equal. ■

EXERCISES

For each of the following exercises state the practical and probabilistic interpretations of the inter-

val that you construct. Identify each component of the interval: point estimate, reliability coeffi-

cient, and standard error. Explain why the reliability coefficients are not the same for all exercises.

6.6.1 Horwitz et al. (A-18) studied 637 persons who were identified by court records from 1967 to 1971

as having experienced abuse or neglect. For a control group, they located 510 subjects who as chil-

dren attended the same elementary school and lived within a five-block radius of those in the

abused/neglected group. In the abused/neglected group, and control group, 114 and 57 subjects,

respectively, had developed antisocial personality disorders over their lifetimes. Construct a 95 per-

cent confidence interval for the difference between the proportions of subjects developing antiso-

cial personality disorders one might expect to find in the populations of subjects from which the

subjects of this study may be presumed to have been drawn.

6.6.2 The objective of a randomized controlled trial by Adab et al. (A-19) was to determine whether pro-

viding women with additional information on the pros and cons of screening for cervical cancer would

.0791, .4171

.2481 ; 2.581.06552

= .0655

sN pNF -pN M = C

1.455921.54412

68

+ 1.207821.79222

255

pNF - pNM = .4559 - .2078 = .2481.

31>68 = .4559 pNM = 53>255 = .2078.

pNF =

188 CHAPTER 6 ESTIMATION

6.7 DETERMINATION OF SAMPLE SIZE FOR ESTIMATING MEANS

increase the willingness to be screened. A treatment group of 138 women received a leaflet on screen-

ing that contained more information (average individual risk for cervical cancer, likelihood of positive

finding, the possibility of false positive/negative results, etc.) than the standard leaflet developed by

the British National Health Service that 136 women in a control group received. In the treatment group,

109 women indicated they wanted to have the screening test for cervical cancer while in the control

group, 120 indicated they wanted the screening test. Construct a 95 percent confidence interval for the

difference in proportions for the two populations represented by these samples.

6.6.3 Spertus et al. (A-20) performed a randomized single blind study for subjects with stable coronary

artery disease. They randomized subjects into two treatment groups. The first group had current

angina medications optimized, and the second group was tapered off existing medications and then

started on long-acting diltiazem at 180 mg/day. The researchers performed several tests to deter-

mine if there were significant differences in the two treatment groups at baseline. One of the char-

acteristics of interest was the difference in the percentages of subjects who had reported a history

of congestive heart failure. In the group where current medications were optimized, 16 of 49 sub-

jects reported a history of congestive heart failure. In the subjects placed on the diltiazem, 12 of

the 51 subjects reported a history of congestive heart failure. State the assumptions that you think

are necessary and construct a 95 percent confidence interval for the difference between the pro-

portions of those reporting congestive heart failure within the two populations from which we pre-

sume these treatment groups to have been selected.

6.6.4 To study the difference in drug therapy adherence among subjects with depression who received usual

care and those who received care in a collaborative care model was the goal of a study conducted

by Finley et al. (A-21). The collaborative care model emphasized the role of clinical pharmacists in

providing drug therapy management and treatment follow-up. Of the 50 subjects receiving usual care,

24 adhered to the prescribed drug regimen, while 50 out of 75 subjects in the collaborative care model

adhered to the drug regimen. Construct a 90 percent confidence interval for the difference in adherence

proportions for the populations of subjects represented by these two samples.

6.7 DETERMINATION OF SAMPLE SIZE FOR ESTIMATING MEANS

The question of how large a sample to take arises early in the planning of any survey or experiment. This is an important question that should not be treated lightly. To take a larger sample than is needed to achieve the desired results is wasteful of resources, whereas very small samples often lead to results that are of no practical use. Let us consider, then, how one may go about determining the sample size that is needed in a given situation. In this section, we present a method for determining the sample size required for estimating a population mean, and in the next section we apply this method to the case of sample size determination when the parameter to be estimated is a population proportion. By straightforward extensions of these methods, sample sizes required for more complicated situations can be determined.

Objectives The objectives in interval estimation are to obtain narrow intervals with high reliability. If we look at the components of a confidence interval, we see that the width of the interval is determined by the magnitude of the quantity

1reliability coefficient2 \* 1standard error of the estimator2

since the total width of the interval is twice this amount. We have learned that this quan-

tity is usually called the precision of the estimate or the margin of error. For a given

standard error, increasing reliability means a larger reliability coefficient. But a larger reliability coefficient for a fixed standard error makes for a wider interval.

On the other hand, if we fix the reliability coefficient, the only way to reduce the width of the interval is to reduce the standard error. Since the standard error is equal to and since is a constant, the only way to obtain a small standard error is to take a large sample. How large a sample? That depends on the size of the population standard deviation, the desired degree of reliability, and the desired interval width.

Let us suppose we want an interval that extends d units on either side of the estimator. We can write

(6.7.1)

If sampling is to be with replacement, from an infinite population, or from a population that is sufficiently large to warrant our ignoring the finite population correction, Equation 6.7.1 becomes

(6.7.2)

which, when solved for n, gives

(6.7.3)

When sampling is without replacement from a small finite population, the finite population correction is required and Equation 6.7.1 becomes

(6.7.4)

which, when solved for n, gives

(6.7.5)

If the finite population correction can be ignored, Equation 6.7.5 reduces to Equation 6.7.3.

**Estimating**

The formulas for sample size require knowledge of but, as has been pointed out, the population variance is, as a rule, unknown. As a result, has to be estimated. The most frequently used sources of estimates for are the following:

1. A pilot or preliminary sample may be drawn from the population, and the variance computed from this sample may be used as an estimate of Observations used in the pilot sample may be counted as part of the final sample, so that n (the computed sample size) (the pilot sample size) (the number of observations needed to satisfy the total sample size requirement).

d = z s

1n

d = 1reliability coefficient2 \* 1standard error of the estimator2

s,

s>1n, s

190 CHAPTER 6 ESTIMATION

EXERCISES 191

2. Estimates of may be available from previous or similar studies.

3. If it is thought that the population from which the sample is to be drawn is approximately normally distributed, one may use the fact that the range is approximately equal to six standard deviations and compute This method requires some knowledge of the smallest and largest value of the variable in the

population.

**EXAMPLE 6.7.1**

A health department nutritionist, wishing to conduct a survey among a population of teenage girls to determine their average daily protein intake (measured in grams), is seeking the advice of a biostatistician relative to the sample size that should be taken.

What procedure does the biostatistician follow in providing assistance to the nutritionist? Before the statistician can be of help to the nutritionist, the latter must provide three items of information: (1) the desired width of the confidence interval, (2) the level of confidence desired, and (3) the magnitude of the population variance.

**Solution**: Let us assume that the nutritionist would like an interval about 10 grams wide; that is, the estimate should be within about 5 grams of the population mean in either direction. In other words, a margin of error of 5 grams is desired. Let us also assume that a confidence coefficient of .95 is decided on and that, from past experience, the nutritionist feels that the population standard deviation is probably about 20 grams. The statistician now has the

necessary information to compute the sample size: and Let us assume that the population of interest is large so that the statistician may ignore the finite population correction and use Equation 6.7.3. On making proper substitutions, the value of n is found to be

The nutritionist is advised to take a sample of size 62. When calculating a sample size by Equation 6.7.3 or Equation 6.7.5, we round up to the next-largest whole number if the calculations yield a number that is not itself an integer. ■

**EXERCISES**

6.7.1 A hospital administrator wishes to estimate the mean weight of babies born in her hospital. How

large a sample of birth records should be taken if she wants a 99 percent confidence interval that

is 1 pound wide? Assume that a reasonable estimate of is 1 pound. What sample size is required

if the confidence coefficient is lowered to .95?

6.7.2 The director of the rabies control section in a city health department wishes to draw a sample from

the department’s records of dog bites reported during the past year in order to estimate the mean

s

= 61.47

n = 11.9622 12022

1522

d = 5.

z = 1.96, s = 20,

s L R>6.

s2

age of persons bitten. He wants a 95 percent confidence interval, he will be satisfied to let

and from previous studies he estimates the population standard deviation to be about 15 years.

How large a sample should be drawn?

6.7.3 A physician would like to know the mean fasting blood glucose value (milligrams per 100 ml) of

patients seen in a diabetes clinic over the past 10 years. Determine the number of records the

physician should examine in order to obtain a 90 percent confidence interval for if the desired

width of the interval is 6 units and a pilot sample yields a variance of 60.

6.7.4 For multiple sclerosis patients we wish to estimate the mean age at which the disease was first

diagnosed. We want a 95 percent confidence interval that is 10 years wide. If the population vari-

ance is 90, how large should our sample be?

6.8 DETERMINATION OF SAMPLE SIZE FOR ESTIMATING PROPORTIONS

The method of sample size determination when a population proportion is to be esti-

mated is essentially the same as that described for estimating a population mean. We

make use of the fact that one-half the desired interval, d, may be set equal to the prod-

uct of the reliability coefficient and the standard error.

Assuming that random sampling and conditions warranting approximate nor-

mality of the distribution of leads to the following formula for n when sampling

is with replacement, when sampling is from an infinite population, or when the sam-

pled population is large enough to make use of the finite population correction

unnecessary,

(6.8.1)

where

If the finite population correction cannot be disregarded, the proper formula for

n is

(6.8.2)

When N is large in comparison to n (that is, the finite population cor-

rection may be ignored, and Equation 6.8.2 reduces to Equation 6.8.1.

Estimating p As we see, both formulas require knowledge of p, the proportion in

the population possessing the characteristic of interest. Since this is the parameter we

are trying to estimate, it, obviously, will be unknown. One solution to this problem is to

take a pilot sample and compute an estimate to be used in place of p in the formula for

n. Sometimes an investigator will have some notion of an upper bound for p that can be

used in the formula. For example, if it is desired to estimate the proportion of some pop-

ulation who have a certain disability, we may feel that the true proportion cannot be

n>N ... .052

n = Nz 2

pq

d2

1N - 12 + z 2

pq

q = 1 - p.

n = z 2

pq

d2

pN

m

d = 2.5,

192 CHAPTER 6 ESTIMATION

EXERCISES 193

greater than, say, .30. We then substitute .30 for p in the formula for n. If it is impossi-

ble to come up with a better estimate, one may set p equal to .5 and solve for n. Since

in the formula yields the maximum value of n, this procedure will give a large

enough sample for the desired reliability and interval width. It may, however, be larger

than needed and result in a more expensive sample than if a better estimate of p had

been available. This procedure should be used only if one is unable to arrive at a better

estimate of p.

EXAMPLE 6.8.1

A survey is being planned to determine what proportion of families in a certain area are

medically indigent. It is believed that the proportion cannot be greater than .35. A 95

percent confidence interval is desired with What size sample of families should

be selected?

Solution: If the finite population correction can be ignored, we have

The necessary sample size, then, is 350. ■

EXERCISES

6.8.1 An epidemiologist wishes to know what proportion of adults living in a large metropolitan area

have subtype ayr hepatitis B virus. Determine the sample size that would be required to estimate

the true proportion to within .03 with 95 percent confidence. In a similar metropolitan area the

proportion of adults with the characteristic is reported to be .20. If data from another metropoli-

tan area were not available and a pilot sample could not be drawn, what sample size would be

required?

6.8.2 A survey is planned to determine what proportion of the high-school students in a metropolitan

school system have regularly smoked marijuana. If no estimate of p is available from previous

studies, a pilot sample cannot be drawn, a confidence coefficient of .95 is desired, and is

to be used, determine the appropriate sample size. What sample size would be required if 99 per-

cent confidence were desired?

6.8.3 A hospital administrator wishes to know what proportion of discharged patients is unhappy with

the care received during hospitalization. How large a sample should be drawn if we let

the confidence coefficient is .95, and no other information is available? How large should the sam-

ple be if p is approximated by .25?

6.8.4 A health planning agency wishes to know, for a certain geographic region, what proportion of

patients admitted to hospitals for the treatment of trauma die in the hospital. A 95 percent confi-

dence interval is desired, the width of the interval must be .06, and the population proportion, from

other evidence, is estimated to be .20. How large a sample is needed?

d = .05,

d = .04

n = 11.9622

1.3521.652

1.0522 = 349.59

d = .05.

p = .5

6.9 CONFIDENCE INTERVAL FOR

THE VARIANCE OF A NORMALLY

DISTRIBUTED POPULATION

**Point Estimation of the Population Variance**

In previous sections it has been suggested that when a population variance is unknown, the sample variance may be used as an estimator. You may have wondered about the quality of this estimator. We have discussed only one criterion of quality—unbiasedness—so let us see if the sample variance is an unbiased estimator of the population variance. To be unbiased,

the average value of the sample variance over all possible samples must be equal to the population variance. That is, the expression must hold. To see if this condition holds for a particular situation, let us refer to the example of constructing a sampling distribution given in Section 5.3. In Table 5.3.1 we have all possible samples of size 2 from the population consisting of the values 6, 8, 10, 12, and 14. It will be recalled that two measures of dispersion for this population were computed as follows:

and

If we compute the sample variance for each of the possible samples shown in Table 5.3.1, we obtain the sample variances shown in Table 6.9.1.

**Sampling with Replacement**

If sampling is with replacement, the expected value of is obtained by taking the mean of all sample variances in Table 6.9.1. When we do this, we have

and we see, for example, that when sampling is with replacement where and s2 = g1xi - m22 s >N. 2 = g1xi - x22

>1n - 12

E1s 2

2 = s2

,

E1s 2

2 = gs 2

i

N n = 0 + 2 + Á + 2 + 0

25 = 200

25 = 8

s2

s 2 = g1xi - x22

>1n - 12

S2 = g1xi - m22

N - 1 s = 10 2 = g1xi - m22

N = 8

E1s 2

2 = s2

194 CHAPTER 6 ESTIMATION

TABLE 6.9.1 Variances Computed from Samples

Shown in Table 5.3.1

Second Draw

6 8 10 12 14

6 0 2 8 18 32

8 2 0 2 8 18

First Draw 10 820 2 8

12 18 8 2 0 2

14 32 18 8 2 0

6.9 CONFIDENCE INTERVAL FOR THE VARIANCE OF A NORMALLY DISTRIBUTED POPULATION 195

**Sampling Without Replacement**

If we consider the case where sampling is without replacement, the expected value of is obtained by taking the mean of all variances above (or below) the principal diagonal. That is,

which, we see, is not equal to but is equal to

These results are examples of general principles, as it can be shown that, in

general,

when sampling is with replacement

when sampling is without replacement

When N is large, and N will be approximately equal and, consequently, and will be approximately equal.

These results justify our use of when computing the sample variance. In passing, let us note that although is an unbiased estimator of is not an unbiased estimator of The bias, however, diminishes rapidly as n increases.

**Interval Estimation of a Population Variance**

With a point estimate available, it is logical to inquire about the construction of a confidence interval for a population variance. Whether we are successful in constructing a confidence interval for will depend on our ability to find an appropriate sampling distribution.

The Chi-Square Distribution Confidence intervals for are usually based on the sampling distribution of If samples of size n are drawn from a normally distributed population, this quantity has a distribution known as the chi-square distribution with degrees of freedom. As we will say more about this distribution in chapter 12, we only say here that it is the distribution that the quantity follows and that it is useful in finding confidence intervals for when the assumption that the population is normally distributed holds true.

Figure 6.9.1 shows chi-square distributions for several values of degrees of freedom. Percentiles of the chi-square distribution are given in Appendix Table F. The column headings give the values of to the left of which lies a proportion of the total area under the curve equal to the subscript of The row labels are the degrees of freedom.

To obtain a percent confidence interval for we first obtain the percent confidence interval for To do this, we select the values of from Appendix Table F in such a way that is to the left of the smaller value and is to the right of the larger value. In other words, the two values of are selected in such a way that is divided equally between the two tails of the distribution. We may

x2 a>2

x a>2 2

1n - 12s 2

>s2 10011 - a2 .

s2 10011 - a2 ,

x2

.

x2

s2

1n - 12s 2

>s2

n - 1

1x2

2

1n - 12s 2

>s2

.

s2

s2

s s. 2

, s

s 2

s 2 = g1xi - x22

>1n - 12

S 2

s2 N - 1

E1s 2

2 = S 2

E1s 2

2 = s2

S2 = g1xi - m22 s >1N - 12. 2

,

E1s 2

2 = gs 2

i

NCn

= 2 + 8 + Á + 2

10 = 100

10 = 10

s2

designate these two values of as and respectively. The percent confidence interval for then, is given by

We now manipulate this expression in such a way that we obtain an expression with alone as the middle term. First, let us divide each term by to get

If we take the reciprocal of this expression, we have

Note that the direction of the inequalities changed when we took the reciprocals. If we reverse the order of the terms, we have

(6.9.1) 1n - 12s2

x2

1-1a>22

6 s2 6 1n - 12s2

x2

a>2

1n - 12s 2

x2

a>2

7 s2 7 1n - 12s 2

x2

1-1a>22

x2

a>2

1n - 12s 2 6

1

s2 6

x2

1-1a>22

1n - 12s 2

1n - 12s

2 s2

x2

a>2 6 1n - 12 s2

s2 6 x2

1-1a>22

1n - 12s 2

>s2

,

x 10011 - a2 2

1-1a>22 x , 2 x a>2

2

196 CHAPTER 6 ESTIMATION

FIGURE 6.9.1 Chi-square distributions for several values of degrees of freedom k.

(Source: Paul G. Hoel and Raymond J. Jessen, Basic Statistics for Business and

Economics, Wiley, 1971. Used with permission.)

6.9 CONFIDENCE INTERVAL FOR THE VARIANCE OF A NORMALLY DISTRIBUTED POPULATION 197

which is the percent confidence interval for If we take the square root

of each term in Expression 6.9.1, we have the following percent confidence

interval for the population standard deviation:

(6.9.2)

**EXAMPLE 6.9.1**

In a study of the effectiveness of a gluten-free diet in first-degree relatives of patients with type I diabetics, Hummel et al. (A-22) placed seven subjects on a gluten-free diet for 12 months. Prior to the diet, they took baseline measurements of several antibodies and autoantibodies, one of which was the diabetes related insulin autoantibody (IAA). The IAA levels were measured by radiobinding assay. The seven subjects had IAA units of

9.7, 12.3, 11.2, 5.1, 24.8, 14.8, 17.7

We wish to estimate from the data in this sample the variance of the IAA units in the population from which the sample was drawn and construct a 95 percent confidence interval for this estimate.

**Solution**: The sample yielded a value of The degrees of freedom are The appropriate values of from Appendix Table F are and Our 95 percent confidence interval for

is

The 95 percent confidence interval for is

We are 95 percent confident that the parameters being estimated are within the specified limits, because we know that in the long run, in repeated sampling, 95 percent of intervals constructed as illustrated would include the respective parameters. ■

**Some Precautions**

Although this method of constructing confidence intervals for is widely used, it is not without its drawbacks. First, the assumption of the normality of the population from which the sample is drawn is crucial, and results may be misleading if the assumption is ignored.

Another difficulty with these intervals results from the fact that the estimator is not in the center of the confidence interval, as is the case with the confidence interval for

This is because the chi-square distribution, unlike the normal, is not symmetric. The prac-

tical implication of this is that the method for the construction of confidence intervals

m.

s2

4.063 6 s 6 13.888

s

16.512 6 s2 6 192.868

6139.7632

14.449

6 s2 6

6139.7632

1.237

s2

x2 x 14.449 a>2 = 1.237. 2

1-1a>22 =

x2 n - 1 = 6.

s 2 = 39.763.

C

1n - 12s 2

x2

1-1a>22

6 s 6 C

1n - 12s 2

x2

a>2

s,

10011 - a2

s2 10011 - a2 .

for which has just been described, does not yield the shortest possible confidence

intervals. Tate and Klett (12) give tables that may be used to overcome this difficulty.

**EXERCISES**

6.9.1 A study by Aizenberg et al. (A-23) examined the efficacy of sildenafil, a potent phosphodiesterase

inhibitor, in the treatment of elderly men with erectile dysfunction induced by antidepressant treat-

ment for major depressive disorder. The ages of the 10 enrollees in the study were

74, 81, 70, 70, 74, 77, 76, 70, 71, 72

Assume that the subjects in this sample constitute a simple random sample drawn from a popula-

tion of similar subjects. Construct a 95 percent confidence interval for the variance of the ages of

subjects in the population.

6.9.2 Borden et al. (A-24) performed experiments on cadaveric knees to test the effectiveness of several

meniscal repair techniques. Specimens were loaded into a servohydraulic device and tension-loaded

to failure. The biomechanical testing was performed by using a slow loading rate to simulate the

stresses that the medial meniscus might be subjected to during early rehabilitation exercises and

activities of daily living. One of the measures is the amount of displacement that occurs. Of the

12 specimens receiving the vertical mattress suture and the FasT-FIX method, the displacement

values measured in millimeters are 16.9, 20.2, 20.1, 15.7, 13.9, 14.9, 18.0, 18.5, 9.2, 18.8, 22.8,

17.5. Construct a 90 percent confidence interval for the variance of the displacement in millime-

ters for a population of subjects receiving these repair techniques.

6.9.3 Forced vital capacity determinations were made on 20 healthy adult males. The sample variance

was 1,000,000. Construct 90 percent confidence intervals for and

6.9.4 In a study of myocardial transit times, appearance transit times were obtained on a sample of

30 patients with coronary artery disease. The sample variance was found to be 1.03. Construct

99 percent confidence intervals for and

6.9.5 A sample of 25 physically and mentally healthy males participated in a sleep experiment in which

the percentage of each participant’s total sleeping time spent in a certain stage of sleep was

recorded. The variance computed from the sample data was 2.25. Construct 95 percent confidence

intervals for and

6.9.6 Hemoglobin determinations were made on 16 animals exposed to a harmful chemical. The follow-

ing observations were recorded: 15.6, 14.8, 14.4, 16.6, 13.8, 14.0, 17.3, 17.4, 18.6, 16.2, 14.7, 15.7,

16.4, 13.9, 14.8, 17.5. Construct 95 percent confidence intervals for and

6.9.7 Twenty air samples taken at the same site over a period of 6 months showed the following amounts

of suspended particulate matter (micrograms per cubic meter of air):

68 22 36 32

42 24 28 38

30 44 28 27

28 43 45 50

79 74 57 21

Consider these measurements to be a random sample from a population of normally distributed

measurements, and construct a 95 percent confidence interval for the population variance.

s s. 2

s s. 2

s s. 2

s s. 2

s2

,

198 CHAPTER 6 ESTIMATION

6.10 CONFIDENCE INTERVAL FOR THE RATIO OF TWO VARIANCES 199

6.10 CONFIDENCE INTERVAL

FOR THE RATIO OF THE VARIANCES

OF TWO NORMALLY DISTRIBUTED

POPULATIONS

It is frequently of interest to compare two variances, and one way to do this is to form their ratio, If two variances are equal, their ratio will be equal to 1. We usually will not know the variances of populations of interest, and, consequently, any comparisons we make will be based on sample variances. In other words, we may wish to estimate the ratio of two population variances. We learned in Section 6.4 that the valid use of the t distribution to construct a confidence interval for the difference between two population means requires that the population variances be equal. The use of the ratio of two population variances for determining equality of variances has been formalized into a statistical test. The distribution of this test provides test values for determining if the ratio exceeds the value 1 to a large enough extent that we may conclude that the variances are not equal. The test is referred to as the F-max Test by Hartley (13) or the Variance Ratio Test by Zar (14). Many computer programs provide some formalized test of the equality of variances so that the assumption of equality of variances associated with many of the tests in the following chapters can be examined. If the confidence interval for the ratio of two population variances includes 1, we conclude that the two population variances may, in fact, be equal. Again, since this is a form of inference, we must rely on some sampling distribution, and this time the distribution of is utilized provided certain assumptions are met. The assumptions are that and are computed from independent samples of size and respectively, drawn from two normally distributed populations. We use to designate the larger of the two sample variances.

**The F Distribution**

If the assumptions are met, follows a distribution known as the F distribution. We defer a more complete discussion of this distribution until chapter 8, but note that this distribution depends on two-degrees-of-freedom values, one corresponding to the value of used in computing and the other corresponding to the value of used in computing These are usually referred to as the numerator degrees of freedom and the denominator degrees of freedom. Figure 6.10.1 shows some F distributions for several numerator and denominator degrees-of-freedom combinations. Appendix Table G contains, for specified combinations of degrees of freedom and values of values to the right of which lies of the area under the curve of F.

**A Confidence Interval for**

To find the percent confidence interval for we begin with the expression

where and are the values from the F table to the left and right of which, respectively, lies of the area under the curve. The middle term of this expression may be rewritten so that the entire expression is

If we divide through by we have

Taking the reciprocals of the three terms gives

and if we reverse the order, we have the following percent confidence inter-

val for

(6.10.1)

**EXAMPLE 6.10.1**

Allen and Gross (A-25) examine toe flexors strength in subjects with plantar fasciitis (pain

from heel spurs, or general heel pain), a common condition in patients with musculoskeletal problems. Inflammation of the plantar fascia is often costly to treat and frustrating for both the patient and the clinician. One of the baseline measurements was the body mass index (BMI). For the 16 women in the study, the standard deviation for BMI was 8.1 and for four men in the study, the standard deviation was 5.9. We wish to construct a 95 percent confidence interval for the ratio of the variances of the two populations from which we presume these samples were drawn.

**Solution**: We have the following information:

We are now ready to obtain our 95 percent confidence interval for

by substituting appropriate values into Expression 6.10.1:

We give this interval the appropriate probabilistic and practical interpretations.

Since the interval .1323 to 7.8221 includes 1, we are able to conclude

that the two population variances may be equal. ■

**Finding and**

At this point we must make a cumbersome, but unavoidable, digression and explain how the values and were obtained. The value of at the intersection of the column headed and the row labeled is 14.25. If we had a more extensive table of the F distribution, finding would be no trouble; we would simply find as we found We would take the value at the intersection of the column headed 15 and the row headed 3. To include every possible percentile of F would make for a very lengthy table. Fortunately, however, there exists a relationship that enables us to compute the lower percentile values from our limited table. The relationship is as follows:

(6.10.2)

We proceed as follows.

Interchange the numerator and denominator degrees of freedom and locate the appropriate value of F. For the problem at hand we locate 4.15, which is at the intersection of the column headed 3 and the row labeled 15. We now take the reciprocal of this value, In summary, the lower confidence limit (LCL) and upper confidence limit (UCL) are as follows:

Alternative procedures for making inferences about the equality of two variances

when the sampled populations are not normally distributed may be found in the book by

Daniel (15).

**Some Precautions** Similar to the discussion in the previous section of constructing confidence intervals for , the assumption of normality of the populations from which the samples are drawn is crucial to obtaining correct intervals for the ratio of variances discussed in this section. Fortunately, most statistical computer programs provide alternatives to the F-ratio, such as Levene’s test, when the underlying distributions cannot be assumed to be normally distributed. Computationally, Levene’s test uses a measure of distance from a sample median instead of a sample mean, hence removing the assumption of normality.

**EXERCISES**

6.10.1 The purpose of a study by Moneim et al. (A-26) was to examine thumb amputations from team

roping at rodeos. The researchers reviewed 16 cases of thumb amputations. Of these, 11 were com-

plete amputations while five were incomplete. The ischemia time is the length of time that insuf-

ficient oxygen is supplied to the amputated thumb. The ischemia times (hours) for 11 subjects

experiencing complete amputations were

4.67, 10.5, 2.0, 3.18, 4.00, 3.5, 3.33, 5.32, 2.0, 4.25, 6.0

For five victims of incomplete thumb amputation, the ischemia times were

3.0, 10.25, 1.5, 5.22, 5.0

Treat the two reported sets of data as sample data from the two populations as described.

Construct a 95 percent confidence interval for the ratio of the two unknown population

variances.

6.10.2 The objective of a study by Horesh et al. (A-27) was to explore the hypothesis that some forms

of suicidal behavior among adolescents are related to anger and impulsivity. The sample consisted

of 65 adolescents admitted to a university-affiliated adolescent psychiatric unit. The researchers

used the Impulsiveness-Control Scale (ICS, A-28) where higher numbers indicate higher degrees

of impulsiveness and scores can range from 0 to 45. The 33 subjects classified as suicidal had an

ICS score standard deviation of 8.4 while the 32 nonsuicidal subjects had a standard deviation of

6.0. Assume that these two groups constitute independent simple random samples from two

s2

UCL = s 2

1

s 2

2

F1-1a>22, df2,df1

LCL = s 2

1

s 2

2

1

F11-a>22,df1,df2

s2

1>s2

2

1>4.15 = .24096.

202 CHAPTER 6 ESTIMATION

6.11 SUMMARY 203

populations of similar subjects. Assume also that the ICS scores in these two populations are nor-

mally distributed. Find the 99 percent confidence interval for the ratio of the two population vari-

ances of scores on the ICS.

6.10.3 Stroke index values were statistically analyzed for two samples of patients suffering from

myocardial infarction. The sample variances were 12 and 10. There were 21 patients in

each sample. Construct the 95 percent confidence interval for the ratio of the two population

variances.

6.10.4 Thirty-two adult asphasics seeking speech therapy were divided equally into two groups. Group 1

received treatment 1, and group 2 received treatment 2. Statistical analysis of the treatment effec-

tiveness scores yielded the following variances: Construct the 90 percent confi-

dence interval for

6.10.5 Sample variances were computed for the tidal volumes (milliliters) of two groups of patients suf-

fering from atrial septal defect. The results and sample sizes were as follows:

Construct the 95 percent confidence interval for the ratio of the two population variances.

6.10.6 Glucose responses to oral glucose were recorded for 11 patients with Huntington’s disease (group 1)

and 13 control subjects (group 2). Statistical analysis of the results yielded the following sample

variances: Construct the 95 percent confidence interval for the ratio of the

two population variances.

6.10.7 Measurements of gastric secretion of hydrochloric acid (milliequivalents per hour) in 16 normal

subjects and 10 subjects with duodenal ulcer yielded the following results:

Normal subjects: 6.3, 2.0, 2.3, 0.5, 1.9, 3.2, 4.1, 4.0, 6.2, 6.1,

3.5, 1.3, 1.7, 4.5, 6.3, 6.2

Ulcer subjects: 13.7, 20.6, 15.9, 28.4, 29.4, 18.4, 21.1, 3.0,

26.2, 13.0

Construct a 95 percent confidence interval for the ratio of the two population variances. What

assumptions must be met for this procedure to be valid?

6.11 SUMMARY

This chapter is concerned with one of the major areas of statistical inference—estimation. Both point estimation and interval estimation are covered. The concepts and methods

involved in the construction of confidence intervals are illustrated for the following

parameters: means, the difference between two means, proportions, the difference between

two proportions, variances, and the ratio of two variances. In addition, we learned in this

chapter how to determine the sample size needed to estimate a population mean and a

population proportion at specified levels of precision.

We learned, also, in this chapter that interval estimates of population parameters

are more desirable than point estimates because statements of confidence can be attached

to interval estimates.

s 2

1 = 105, s 2

2 = 148.

n 2 = 41, s 2

2 = 20,000

n1 = 31, s 2

1 = 35,000

s2

2 >s2

1.

s 2

1 = 8, s 2

2 = 15.

SUMMARY OF FORMULAS FOR CHAPTER 6

Formula Number Name Formula

6.2.1 Expression of estimator (reliability coefficient)

an interval estimate ( (standard error of the estimator)

6.2.2 Interval estimate

for when is

known

6.3.1 t-transformation

6.3.2 Interval estimate

for when is

unknown

6.4.1 Interval estimate

for the difference

between two

population means

when and are

known

6.4.2 Pooled variance

estimate

6.4.3 Standard error

of estimate

6.4.4 Interval estimate

for the difference

between two

population means

when is unknown

6.4.5 Cochran’s

correction for

reliability coefficient

when variances

are not equal

6.4.6 Interval estimate

using Cochran’s

correction for t

6.5.1 Interval estimate

for a population

proportion

pN ; z 11-a>221pN11 - pN2>n

1x 1 - x 22 ; t¿11-a>22

A

s1

2

n1 + s2

2

n2

t¿11-a>22 = w1t1 + w2t2

w1 + w2

s

1x 1 - x 22 ; t11-a>22

D

sp

2

n1

+ sp

2

n2

s1x1-x22 = D

sp

2

n1

+ sp

2

n2

sp

2 = 1n1 - 12s1

2 + 1n2 - 12s2

2

n1 + n2 - 2

s1 s2

1x 1 - x 22 ; z 11-a>22

A

s1

2

n1

+ s2

2

n2

m s x ; t11-a>22 = s

2n

t = x - m

s> 1n

m s

x ; z 11-a>22sx

;

204 CHAPTER 6 ESTIMATION

SUMMARY OF FORMULAS FOR CHAPTER 6 205

(Continued)

6.6.1 Interval estimate

for the difference

between two

population

proportions

6.7.1–6.7.3 Sample size d " (reliability coefficient) ( (standard error)

determination

when sampling

with replacement

6.7.4–6.7.5 Sample size

determination when

sampling without

replacement

6.8.1 Sample size

determination

for proportions

when sampling

with replacement

6.8.2 Sample size

determination for

proportions when

sampling without

replacement

6.9.1 Interval estimate

for

6.9.2 Interval estimate

for

6.10.1 Interval estimate

for the ratio of

two variances

6.10.2 Relationship

among F ratios

Symbol Key • " Type 1 error rate

• " Chi-square distribution

• d " error component of interval estimate

x2

a

Fa,df1,df2 = 1

F1-a,df2,df1

s1

2

>s2

2

F1-1a>22

6 s1

2

s2

2 6 s1

2

>s2

2

Fa>2

s C

1n - 12s2

x2

1-1a>22

6 s 6 C

1n - 12s2

x2

1a>22

s2 1n - 12s2

x2

1-1a>22

6 s2 6

1n - 12s2

x2

1a>22

n = Nz2

s2

d2

1N - 12 + z2

s2

n = z

2

pq

d2

1p1N - pN22 ; z 11-a>22

A

p1N 11 - p1N 2

n1

+

p2N 11 - p2N 2

n2

n = N z2

s2

d2

1N - 12 + z2

s2

‹

d = z s

1nA

N - n

N - 1

n = z2

s2

d2

‹

d = z s

1n

206 CHAPTER 6 ESTIMATION

• df " degrees of freedom

• F " F-distribution

• " mean of population

• n " sample size

• p " proportion for population

• q " (1 & p)

• " estimated proportion for sample

• " population variance

• " population standard deviation

• " standard error

• s " standard deviation of sample

• sp " pooled standard deviation

• t " Student’s t-transformation

• t' " Cochran’s correction to t

• " mean of sample

• z " standard normal distribution

REVIEW QUESTIONS AND EXERCISES

1. What is statistical inference?

2. Why is estimation an important type of inference?

3. What is a point estimate?

4. Explain the meaning of unbiasedness.

5. Define the following:

(a) Reliability coefficient (c) Precision (e) Estimator

(b) Confidence coefficient (d) Standard error (f) Margin of error

6. Give the general formula for a confidence interval.

7. State the probabilistic and practical interpretations of a confidence interval.

8. Of what use is the central limit theorem in estimation?

9. Describe the t distribution.

10. What are the assumptions underlying the use of the t distribution in estimating a single popula-

tion mean?

11. What is the finite population correction? When can it be ignored?

12. What are the assumptions underlying the use of the t distribution in estimating the difference

between two population means?

13. Arterial blood gas analyses performed on a sample of 15 physically active adult males yielded the

following resting values:

75, 80, 80, 74, 84, 78, 89, 72, 83, 76, 75, 87, 78, 79, 88

PaO2

x

sx

s

s2

pN

m

Compute the 95 percent confidence interval for the mean of the population.

14. What proportion of asthma patients are allergic to house dust? In a sample of 140, 35 percent

had positive skin reactions. Construct the 95 percent confidence interval for the population

proportion.

15. An industrial hygiene survey was conducted in a large metropolitan area. Of 70 manufacturing

plants of a certain type visited, 21 received a “poor” rating with respect to absence of safety haz-

ards. Construct a 95 percent confidence interval for the population proportion deserving a “poor”

rating.

16. Refer to the previous problem. How large a sample would be required to estimate the population

proportion to within .05 with 95 percent confidence (.30 is the best available estimate of p):

(a) If the finite population correction can be ignored?

(b) If the finite population correction is not ignored and N 1500?

17. In a dental survey conducted by a county dental health team, 500 adults were asked to give the

reason for their last visit to a dentist. Of the 220 who had less than a high-school education, 44

said they went for preventative reasons. Of the remaining 280, who had a high-school educa-

tion or better, 150 stated that they went for preventative reasons. Construct a 95 percent confi-

dence interval for the difference between the two population proportions.

18. A breast cancer research team collected the following data on tumor size:

Type of Tumor n s

A 21 3.85 cm 1.95 cm

B 16 2.80 cm 1.70 cm

Construct a 95 percent confidence interval for the difference between population means.

19. A certain drug was found to be effective in the treatment of pulmonary disease in 180 of 200

cases treated. Construct the 90 percent confidence interval for the population proportion.

20. Seventy patients with stasis ulcers of the leg were randomly divided into two equal groups. Each

group received a different treatment for edema. At the end of the experiment, treatment effective-

ness was measured in terms of reduction in leg volume as determined by water displacement. The

means and standard deviations for the two groups were as follows:

Group (Treatment) s

A 95 cc 25

B 125 cc 30

Construct a 95 percent confidence interval for the difference in population means.

21. What is the average serum bilirubin level of patients admitted to a hospital for treatment of hep-

atitis? A sample of 10 patients yielded the following results:

20.5, 14.8, 21.3, 12.7, 15.2, 26.6, 23.4, 22.9, 15.7, 19.2

Construct a 95 percent confidence interval for the population mean.

x

x

=

REVIEW QUESTIONS AND EXERCISES 207

22. Determinations of saliva pH levels were made in two independent random samples of seventh-

grade schoolchildren. Sample A children were caries-free while sample B children had a high

incidence of caries. The results were as follows:

A: 7.14, 7.11, 7.61, 7.98, 7.21, 7.16, 7.89 B: 7.36, 7.04, 7.19, 7.41, 7.10, 7.15, 7.36,

7.24, 7.86, 7.47, 7.82, 7.37, 7.66, 7.62, 7.65 7.57, 7.64, 7.00, 7.25, 7.19

Construct a 90 percent confidence interval for the difference between the population means.

Assume that the population variances are equal.

23. Drug A was prescribed for a random sample of 12 patients complaining of insomnia. An independ-

ent random sample of 16 patients with the same complaint received drug B. The number of hours of

sleep experienced during the second night after treatment began were as follows:

A: 3.5, 5.7, 3.4, 6.9, 17.8, 3.8, 3.0, 6.4, 6.8, 3.6, 6.9, 5.7

B: 4.5, 11.7, 10.8, 4.5, 6.3, 3.8, 6.2, 6.6, 7.1, 6.4, 4.5,

5.1, 3.2, 4.7, 4.5, 3.0

Construct a 95 percent confidence interval for the difference between the population means.

Assume that the population variances are equal.

24. The objective of a study by Crane et al. (A-29) was to examine the efficacy, safety, and maternal

satisfaction of (a) oral misoprostol and (b) intravenous oxytocin for labor induction in women with

premature rupture of membranes at term. Researchers randomly assigned women to the two treat-

ments. For the 52 women who received oral misoprostol, the mean time in minutes to active labor

was 358 minutes with a standard deviation of 308 minutes. For the 53 women taking oxytocin,

the mean time was 483 minutes with a standard deviation of 144 minutes. Construct a 99 percent

confidence interval for the difference in mean time to active labor for these two different medica-

tions. What assumptions must be made about the reported data? Describe the population about

which an inference can be made.

25. Over a 2-year period, 34 European women with previous gestational diabetes were retrospec-

tively recruited from West London antenatal databases for a study conducted by Kousta et al.

(A-30). One of the measurements for these women was the fasting nonesterified fatty acids con-

centration (NEFA) measured in In the sample of 34 women, the mean NEFA level was

435 with a sample standard deviation of 215.0. Construct a 95 percent confidence interval for

the mean fasting NEFA level for a population of women with gestational diabetes. State all

necessary assumptions about the reported data and subjects.

26. Scheid et al. (A-31) questioned 387 women receiving free bone mineral density screening. The

questions focused on past smoking history. Subjects undergoing hormone replacement therapy

(HRT), and subjects not undergoing HRT, were asked if they had ever been a regular smoker. In

the HRT group, 29.3 percent of 220 women stated that they were at some point in their life a reg-

ular smoker. In the non–HRT group, 17.3 percent of 106 women responded positively to being at

some point in their life a regular smoker. (Sixty-one women chose not to answer the question.)

Construct a 95 percent confidence interval for the difference in smoking percentages for the two

populations of women represented by the subjects in the study. What assumptions about the data

are necessary?

27. The purpose of a study by Elliott et al. (A-32) was to assess the prevalence of vitamin D defi-

ciency in women living in nursing homes. The sample consisted of 39 women in a 120-bed skilled

nursing facility. Women older than 65 years of age who were long-term residents were invited

to participate if they had no diagnosis of terminal cancer or metastatic disease. In the sample,

mmol/L.

208 CHAPTER 6 ESTIMATION

23 women had 25-hydroxyvitamin D levels of 20 ng/ml or less. Construct a 95 percent confidence

interval for the percent of women with vitamin D deficiency in the population presumed to be rep-

resented by this sample.

28. In a study of the role of dietary fats in the etiology of ischemic heart disease the subjects were

60 males between 40 and 60 years of age who had recently had a myocardial infarction and 50

apparently healthy males from the same age group and social class. One variable of interest in

the study was the proportion of linoleic acid (L.A.) in the subjects’ plasma triglyceride fatty acids.

The data on this variable were as follows:

Subjects with Myocardial Infarction

Subject L.A. Subject L.A. Subject L.A. Subject L.A.

1 18.0 2 17.6 3 9.6 4 5.5

5 16.8 6 12.9 7 14.0 8 8.0

9 8.9 10 15.0 11 9.3 12 5.8

13 8.3 14 4.8 15 6.9 16 18.3

17 24.0 18 16.8 19 12.1 20 12.9

21 16.9 22 15.1 23 6.1 24 16.6

25 8.7 26 15.6 27 12.3 28 14.9

29 16.9 30 5.7 31 14.3 32 14.1

33 14.1 34 15.1 35 10.6 36 13.6

37 16.4 38 10.7 39 18.1 40 14.3

41 6.9 42 6.5 43 17.7 44 13.4

45 15.6 46 10.9 47 13.0 48 10.6

49 7.9 50 2.8 51 15.2 52 22.3

53 9.7 54 15.2 55 10.1 56 11.5

57 15.4 58 17.8 59 12.6 60 7.2

Healthy Subjects

Subject L.A. Subject L.A. Subject L.A. Subject L.A.

1 17.1 2 22.9 3 10.4 4 30.9

5 32.7 6 9.1 7 20.1 8 19.2

9 18.9 10 20.3 11 35.6 12 17.2

13 5.8 14 15.2 15 22.2 16 21.2

17 19.3 18 25.6 19 42.4 20 5.9

21 29.6 22 18.2 23 21.7 24 29.7

25 12.4 26 15.4 27 21.7 28 19.3

29 16.4 30 23.1 31 19.0 32 12.9

33 18.5 34 27.6 35 25.0 36 20.0

37 51.7 38 20.5 39 25.9 40 24.6

41 22.4 42 27.1 43 11.1 44 32.7

45 13.2 46 22.1 47 13.5 48 5.3

49 29.0 50 20.2

Construct the 95 percent confidence interval for the difference between population means. What do

these data suggest about the levels of linoleic acid in the two sampled populations?

REVIEW QUESTIONS AND EXERCISES 209

29. The purpose of a study by Tahmassebi and Curzon (A-33) was to compare the mean salivary flow

rate among subjects with cerebral palsy and among subjects in a control group. Each group had 10

subjects. The following table gives the mean flow rate in ml/minute as well as the standard error.

Group Sample Size Mean ml/minute Standard Error

Cerebral palsy 10 0.220 0.0582

Control 10 0.334 0.1641

Source: J. F. Tahmassebi and M. E. J. Curzon, “The Cause of Drooling in Children

with Cerebral Palsy—Hypersalivation or Swallowing Defect?” International Journal of

Paediatric Dentistry, 13 (2003), 106–111.

Construct the 90 percent confidence interval for the difference in mean salivary flow rate for the

two populations of subjects represented by the sample data. State the assumptions necessary for

this to be a valid confidence interval.

30. Culligan et al. (A-34) compared the long-term results of two treatments: (a) a modified Burch pro-

cedure, and (b) a sling procedure for stress incontinence with a low-pressure urethra. Thirty-six

women took part in the study with 19 in the Burch treatment group and 17 in the sling procedure

treatment group. One of the outcome measures at three months post-surgery was maximum ure-

thral closure pressure (cm In the Burch group the mean and standard deviation were 16.4

and 8.2 cm, respectively. In the sling group, the mean and standard deviation were 39.8 and 23.0,

respectively. Construct the 99 percent confidence interval for the difference in mean maximum ure-

thral closure pressure for the two populations represented by these subjects. State all necessary

assumptions.

31. In general, narrow confidence intervals are preferred over wide ones. We can make an interval nar-

row by using a small confidence coefficient. For a given set of other conditions, what happens to

the level of confidence when we use a small confidence coefficient? What would happen to the

interval width and the level of confidence if we were to use a confidence coefficient of zero?

32. In general, a high level of confidence is preferred over a low level of confidence. For a given set

of other conditions, suppose we set our level of confidence at 100 percent. What would be the

effect of such a choice on the width of the interval?

33. The subjects of a study by Borland et al. (A-35) were children in acute pain. Thirty-two children who

presented at an emergency room were enrolled in the study. Each child used the visual analogue scale

to rate pain on a scale from 0 to 100 mm. The mean pain score was 61.3 mm with a 95 percent con-

fidence interval of 53.2 mm–69.4 mm. Which would be the appropriate reliability factor for the inter-

val, z or t? Justify your choice. What is the precision of the estimate? The margin of error?

34. Does delirium increase hospital stay? That was the research question investigated by McCusker

et al. (A-36). The researchers sampled 204 patients with prevalent delirium and 118 without delir-

ium. The conclusion of the study was that patients with prevalent delirium did not have a higher

mean length of stay compared to those without delirium. What was the target population? The

sampled population?

35. Assessing driving self-restriction in relation to vision performance was the objective of a study

by West et al. (A-37). The researchers studied 629 current drivers ages 55 and older for 2 years.

The variables of interest were driving behavior, health, physical function, and vision function.

The subjects were part of a larger vision study at the Smith-Kettlewell Eye Research Institute.

A conclusion of the study was that older adults with early changes in spatial vision function

H2O2.

210 CHAPTER 6 ESTIMATION

and depth perception appear to recognize their limitations and restrict their driving. What was

the target population? The sampled population?

36. In a pilot study conducted by Ayouba et al. (A-38), researchers studied 123 children born of HIV-1-

infected mothers in Yaoundé, Cameroon. Counseled and consenting pregnant women were given a

single dose of nevirapine at the onset of labor. Babies were given a syrup containing nevirapine within

the first 72 hours of life. The researchers found that 87 percent of the children were considered not

infected at 6–8 weeks of age. What is the target population? What is the sampled population?

37. Refer to Exercise 2.3.11. Construct a 95 percent confidence interval for the population mean S/R

ratio. Should you use t or z as the reliability coefficient? Why? Describe the population about

which inferences based on this study may be made.

38. Refer to Exercise 2.3.12. Construct a 90 percent confidence interval for the population mean height.

Should you use t or z as the reliability coefficient? Why? Describe the population about which

inferences based on this study may be made.

Exercises for Use with Large Data Sets Available on the Following Website:

www.wiley.com/college/daniel

1. Refer to North Carolina Birth Registry Data NCBIRTH800 with 800 observations (see Large Data

Exercise 1 in Chapter 2). Calculate 95 percent confidence intervals for the following:

(a) the percentage of male children

(b) the mean age of a mother giving birth

(c) the mean weight gained during pregnancy

(d) the percentage of mothers admitting to smoking during pregnancy

(e) the difference in the average weight gained between smoking and nonsmoking mothers

(f) the difference in the average birth weight in grams between married and nonmarried mothers

(g) the difference in the percentage of low birth weight babies between married and nonmarried

mothers

2. Refer to the serum cholesterol levels for 1000 subjects (CHOLEST). Select a simple random sam-

ple of size 15 from this population and construct a 95 percent confidence interval for the popula-

tion mean. Compare your results with those of your classmates. What assumptions are necessary

for your estimation procedure to be valid?

3. Refer to the serum cholesterol levels for 1000 subjects (CHOLEST). Select a simple random sam-

ple of size 50 from the population and construct a 95 percent confidence interval for the proportion

of subjects in the population who have readings greater than 225. Compare your results with those

of your classmates.

4. Refer to the weights of 1200 babies born in a community hospital (BABY WGTS). Draw a sim-

ple random sample of size 20 from this population and construct a 95 percent confidence interval

for the population mean. Compare your results with those of your classmates. What assumptions

are necessary for your estimation procedure to be valid?

5. Refer to the weights of 1200 babies born in a community hospital (BABY WGTS). Draw a simple ran-

dom sample of size 35 from the population and construct a 95 percent confidence interval for the pop-

ulation mean. Compare this interval with the one constructed in Exercise 4.

6. Refer to the heights of 1000 twelve-year-old boys (BOY HGTS). Select a simple random sample

of size 15 from this population and construct a 99 percent confidence interval for the population

mean. What assumptions are necessary for this procedure to be valid?

REVIEW QUESTIONS AND EXERCISES 211

7. Refer to the heights of 1000 twelve-year-old boys (BOY HGTS). Select a simple random sample

of size 35 from the population and construct a 99 percent confidence interval for the population

mean. Compare this interval with the one constructed in Exercise 5.

REFERENCES

Methodology References

1. JOHN A. RICE, Mathematical Statistics and Data Analysis, 2nd ed., Duxbury, Belmont, CA, 1988.

2. W. S. GOSSET (“Student”), “The Probable Error of a Mean,” Biometrika, 6 (1908), 1–25.

3. W. V. BEHRENS, “Ein Beitrag zu Fehlerberechnung bei wenige Beobachtungen,” Landwirtsschaftliche

Jahrbücher, 68 (1929), 807–837.

4. R. A. FISHER, “The Comparison of Samples with Possibly Unequal Variances,” Annals of Eugenics, 9 (1939),

174–180.

5. R. A. FISHER, “The Asymptotic Approach to Behrens’ Integral with Further Tables for the d Test of Signifi-

cance,” Annals of Eugenics, 11 (1941), 141–172.

6. J. NEYMAN, “Fiducial Argument and the Theory of Confidence Intervals,” Biometrika, 32 (1941), 128–150.

7. H. SCHEFFÉ, “On Solutions of the Behrens-Fisher Problem Based on the t-Distribution,” Annals of Mathemat-

ical Statistics, 14 (1943), 35–44.

8. H. SCHEFFÉ, “A Note on the Behrens-Fisher Problem,” Annals of Mathematical Statistics, 15 (1944), 430–432.

9. B. L. WELCH, “The Significance of the Difference Between Two Means When the Population Variances Are

Unequal,” Biometrika, 29 (1937), 350–361.

10. B. L. WELCH, “The Generalization of ‘Student’s Problem When Several Different Population Variances Are

Involved,” Biometrika, 34 (1947), 28–35.

11. WILLIAM G. COCHRAN, “Approximate Significance Levels of the Behrens-Fisher Test,” Biometrics, 20 (1964),

191–195.

12. R. F. TATE and G. W. KLETT, “Optimal Confidence Intervals for the Variance of a Normal Distribution,”

Journal of the American Statistical Association, 54 (1959), 674–682.

13. H. O. HARTLEY, “The Maximum F-Ratio as a Short Cut Test for Heterogeneity of Variances,” Biometrika, 37

(1950), 308–312.

14. J. H. ZAR, Biostatistical Analysis, 4th ed., Prentice-Hall, Upper Saddle River, NJ, 1999.

15. WAYNE W. DANIEL, Applied Nonparametric Statistics, 2nd ed., PWS-KENT, Boston, 1989.

Applications References

A-1. NICOLA MAFFULLI, CHERLY TALLON, JASON WONG, KIM PENG, and ROBERT BLEAKNEY, “Early Weightbear-

ing and Ankle Mobilization after Open Repair of Acute Midsubstance Tears of the Achilles Tendon,” Ameri-

can Journal of Sports Medicine, 31 (2003), 692–700.

A-2. PAUL J. REBER, LUCY A. MARTINEZ, and SANDRA WEINTRAUB, “Artificial Grammar Learning in Alzheimer’s

Disease,” Cognitive, Affective, and Behavioral Neuroscience, 3 (2003), 145–153.

A-3. CHRISTOPHE PEDROLETTI, MARIEANN HÖGMAN, PEKKA MERILÄINEN, LENNART S. NORDVALL, GUNILLA

HEDLIN, and KJELL ALVING, “Nitric Oxide Airway Diffusing Capacity and Mucosal Concentration in Asth-

matic Schoolchildren,” Pediatric Research, 54 (2003), 496–501.

A-4. BRUCE D. BEYNNON, BRADEN C. FLEMING, DAVID L. CHURCHILL, and DANIEL BROWN, “The Effect of Ante-

rior Cruciate Ligament Deficiency and Functional Bracing on Translation of the Tibia Relative to the Femur

during Nonweightbearing and Weightbearing,” American Journal of Sports Medicine, 31 (2003), 99–105.

A-5. LORRAINE DUGOFF, MAURITHA R. EVERETT, LOUIS VONTVER, and GWYN E. BARLEY, “Evaluation of Pelvic

and Breast Examination Skills of Interns in Obstetrics and Gynecology and Internal Medicine,” American Jour-

nal of Obstetrics and Gynecology, 189 (2003), 655–658.

A-6. DIONNE MAYHEW, KAREN M. PERRIN, and WENDY STRUCHEN, “An Analysis of a Healthy Start Smoking

Cessation Program,” American Journal of Health Studies, 17 (2002), 186–190.

A-7. ERIC GRANHOLM, ROBERT ANTHENELLI, RITA MONTEIRO, JOHN SEVCIK, and MARILYN STOLER, “Brief Inte-

grated Outpatient Dual-Diagnosis Treatment Reduces Psychiatric Hospitalizations,” American Journal on

Addictions, 12 (2003), 306–313.

212 CHAPTER 6 ESTIMATION

A-8. SILVIA IANNELO, ANTONINA CAVALERI, PAOLINA MILAZZO, SANTI CANTARELLA, and FRANCESCO BELFIORE,

“Low Fasting Serum Triglyceride Level as a Precocious Marker of Autoimmune Disorders,” Medscape Gen-

eral Medicine, 5(3)(2003).

A-9. EVELYN C. Y. CHAN, SALLY W. VERNON, FREDERICK T. O’DONNELL, CHUL AHN, ANTHONY GRESINGER, and

DONNIE W. AGA, “Informed Consent for Cancer Screening with Prostate-Specific Antigen: How Well Are Men

Getting the Message?” American Journal of Public Health, 93 (2003), 779–785.

A-10. RONALD F. VAN VOLLENHOVEN, SOFIA ERNESTAM, ANDERS HARJU, JOHAN BRATT, and LARS KLARESKOG,

“Etanercept Versus Etanercept Plus Methotrexate: A Registry-Based Study Suggesting that the Combination Is

Clinically More Effacious,” Arthritis Research and Therapy, 5 (2003), R347–R351.

A-11. SATOSHI NOZAWA, KATSUJI SHIMIZU, KEI MIYAMOTO, and MIZUO TANAKA, “Repair of Pars Interarticularis

Defect by Segmental Wire Fixation in Young Athletes with Spondylolysis,” American Journal of Sports Med-

icine, 31 (2003), 359–364.

A-12. MORI J. KRANTZ, ILANA B. KUTINSKY, ALASTAIR D. ROBERTSON, and PHILIP S. MEHLER, “Dose-Related

Effects of Methadone on QT Prolongation in a Series of Patients with Torsade de Pointes,” Pharmacotherapy,

23 (2003), 802–805.

A-13. SUSANNAH FOX and DEBORAH FELLOWS, “Health Searches and Email Have Become More Commonplace, But

There Is Room for Improvement in Searches and Overall Internet Access,” Pew Internet and American Life

Project, www.pewinternet.org/PPF/r/95/report\_display.asp.

A-14. CARLOS M. LUNA, DANIEL BLANZACO, MICHAEL S. NIDERMAN, WALTER MATARUCCO, NATALIO C. BAREDES,

PABLO DESMERY, FERNANDO PALIZAS, GUILLERMO MENGA, FERNANDO RIOS, and CARLOS APEZTEGUIA, “Res-

olution of Ventilator-Associated Pneumonia: Prospective Evaluation of the Clinical Pulmonary Infection Score

as an Early Clinical Predictor of Outcome,” Critical Care Medicine, 31(3)(2003), 676–682.

A-15. AREND F. L. SCHINKEL, JEROEN J. BAX, ERIC BOERSMA, ABDOU ELHENDY, ELENI C. VOURVOURI, JOS R. T.

C. ROELANDT, and DON POLDERMANS, “Assessment of Residual Myocardial Viability in Regions with Chronic

Electrocardiographic Q-Wave Infarction,” American Heart Journal, 144 (2002), 865–869.

A-16. FRIEDERIKE VON ZUR MUHLEN, WEILUM QUAN, DAVID J. D’AGATE, and TODD J. COHEN, “A Study of Carotid

Sinus Massage and Head-Up Tilt Table Testing in Patients with Syncope and Near-Syncope,” Journal of Inva-

sive Cardiology, 14 (2002), 477–482.

A-17. DANIEL F. CONNOR, RONALD J. STEINGARD, JENNIFER J. ANDERSON, and RICHARD H. MELLONI, Jr., “Gender

Differences in Reactive and Proactive Aggression,” Child Psychiatry and Human Development, 33 (2003),

279–294.

A-18. ALLAN V. HORWITZ, CATHY SPATZ WIDOM, JULIE MCLAUGHLIN, and HELENE RASKIN WHITE, “The Impact

of Childhood Abuse and Neglect on Adult Mental Health: A Prospective Study,” Journal of Health and Social

Behavior, 42 (2001), 184–201.

A-19. P. ADAB, T. MARSHALL, A. ROUSE, B. RANDHAWA, H. SANGHA, and N. BHANGOO, “Randomised Controlled

Trial of the Effect of Evidence Based Information on Women’s Willingness to Participate in Cervical Cancer

Screening,” Journal of Epidemiology and Community Health, 57 (2003), 589–593.

A-20. JOHN A. SPERTUS, TIM KEWHURST, CYNTHIA M. DOUGHERTY, and PAUL NICHOL, “Testing the Effectiveness

of Converting Patients to Long-Acting Antianginal Medications: The Quality of Life in Angina Research,”

American Heart Journal, 141 (2001), 550–558.

A-21. PATRICK R. FINLEY, HEIDI R. RENS, JOAN T. PONT, SUSAN L. GESS, CLIFTON LOUIE, SCOTT A. BULL, JANELLE

Y. LEE, and LISA A. BERO, “Impact of a Collaborative Care Model on Depression in a Primary Care Setting:

A Randomized, Controlled Trial,” Pharmacotherapy, 23(9)(2003), 1175–1185.

A-22. MICHAEL HUMMEL, EZIO BoNIFACIO, HEIKE E. NASERKE, and ANETTE G. ZIEGLER, “Elimination of Dietary

Gluten Does Not Reduce Titers of Type 1 Diabetes-Associated Autoantibodies in High-Risk Subjects,” Diabetes

Care, 25 (2002), 1111–1116.

A-23. DOV AIZENBERG, ABRAHAM WEIZMAN, and YORAM BARAK, “Sildenafil for Selective Serotonin Reuptake

Inhibitor-Induced Erectile Dysfunction in Elderly Male Depressed Patients,” Journal of Sex and Marital

Therapy, 29 (2003), 297–303.

A-24. PETER BORDEN, JOHN NYLAND, DAVID N. M. CABORN, and DAVID PIENKOWSKI, “Biomechanical Comparison

of the FasT-Fix Meniscal Repair Suture System with Vertical Mattress Sutures and Meniscus Arrows,” Amer-

ican Journal of Sports Medicine, 31 (2003), 374–378.

A-25. RACHEL H. ALLEN and MICHAEL T. GROSS, “Toe Flexors Strength and Passive Extension Range of Motion of

the First Metatarsophalangeal Joint in Individuals with Plantar Fasciitis,” Journal of Orthopaedic and Sports

Physical Therapy, 33 (2003), 468–477.

REFERENCES 213

A-26. MOHEB S. MONEIM, KEIKHOSROW FIROOZBAKHSH, DOMINIC GROSS, STEVEN D. YOUNG, and GEORGE OMER,

“Thumb Amputations from Team Roping,” American Journal of Sports Medicine, 31 (2003), 728–735.

A-27. NETTA HORESH, ISRAEL ORBACH, DORON GOTHELF, MEIR EFRATI, and ALAN APTER, “Comparison of the Sui-

cidal Behavior of Adolescent Inpatients with Borderline Personality Disorder and Major Depression,” Journal

of Nervous and Mental Disease, 191 (2003), 582–588.

A-28. R. PLUTCHIK, and H. VAN PRAAG, “The Measurement of Suicidality, Aggressivity and Impulsivity,” Clinical

Neuropharmacology, 9(suppl)(1989), 380–382.

A-29. JOAN M. G. CRANE, TINA DELANEY, and DONNA HUTCHENS, “Oral Misoprostol for Premature Rupture of

Membranes at Term,” American Journal of Obstetrics and Gynecology, 189 (2003), 720–724.

A-30. ELENI KOUSTA, NATASHA J. LAWRENCE, IAN F. GODSLAND, ANNA PENNY, VICTOR ANYAOKU, BARBARA A.

MILLAUER, ESTER CELA, DESMOND G. JOHNSTON, STEPHEN ROBINSON, and MARK I. MCCARTHY, “Insulin

Resistance and Beta-Cell Dysfunction in Normoglycaemic European Women with a History of Gestational Dia-

betes,” Clinical Endocrinology, 59 (2003), 289–297.

A-31. DEWEY C. SCHEID, MARIO T. COLEMAN, and ROBERT M. HAMM, “Do Perceptions of Risk and Quality of Life

Affect Use of Hormone Replacement Therapy by Postmenopausal Women?” Journal of the American Board

of Family Practice, 16 (2003), 270–277.

A-32. MARY E. ELLIOTT, NEIL C. BINKLEY, MOLLY CARNES, DAVID R. ZIMMERMAN, KIM PETERSEN, KATHY KNAPP,

JESSICA J. BEHLKE, NANCY AHMANN, and MARY A. KIESER, “Fracture Risks for Women in Long-Term Care:

High Prevalence of Calcaneal Osteoporosis and Hypovitaminosis D,” Pharmacotherapy, 23 (2003), 702–710.

A-33. J. F. TAHMASSEBI and M. E. J. CURZON, “The Cause of Drooling in Children with Cerebral Palsy—

Hypersalivation or Swallowing Defect?” International Journal of Paediatric Dentistry, 13 (2003), 106–111.

A-34. PATRICK J. CULLIGAN, ROGER P. GOLDBERG, and PETER K. SAND, “A Randomized Controlled Trial Compar-

ing a Modified Burch Procedure and a Suburethral Sling: Long-Term Follow-Up,” International Urogynecol-

ogy Journal, 14 (2003), 229–233.

A-35. MEREDITH L. BORLAND, IAN JACOBS, and GARY GEELHOE, “Intranasal Fentanyl Reduces Acute Pain in Chil-

dren in the Emergency Department: A Safety and Efficacy Study,” Emergency Medicine, 14 (2002), 275–280.

A-36. JANE MCCUSKER, MARTIN G. COLE, NANDINI DENDUKURI, and ERIC BELZILE, “Does Delirium Increase Hos-

pital Stay?” Journal of the American Geriatrics Society, 51 (2003), 1539–1546.

A-37. CATHERINE G. WEST, GINNY GILDENGORIN, GUNILLA HAEGERSTROM PORTNOY, LORI A. LOTT, MARILYN E.

SCHNECK, and JOHN A. BRABYN, “Vision and Driving Self-Restriction in Older Adults,” Journal of the Amer-

ican Geriatrics Society, 51 (2003), 1348–1355.

A-38. AHIDJO AYOUBA, GILBERT TENE, PATRICK CUNIN, YACOUBA FOUPOUAPOUOGNIGNI, ELISABETH MENU,

ANFUMBOM KFUTWAH, JOCELYN THONNON, GABRIELLA SCARLATTI, MARCEL MONNY-LOBÉ, NICOLE ETEKI,

CHARLES KOUANFACK, MICHÈLE TARDY, ROBERT LEKE, MAURICE NKAM, ANNE E. NLEND, FRANÇOISE

BARRÉ-SINOUSSI, PAUL M. V. MARTIN, and ERIC NERRIENET, “Low Rate of Mother-to-Child Transmission of

HIV-1 After Nevirapine Intervention in a Pilot Public Health Program in Yaoundé, Cameroon,” Journal of

Acquired Immune Deficiency Syndrome, 34 (2003), 274–280.

214 CHAPTER 6 ESTIMATION

215

CHAPTER OVERVIEW

This chapter covers hypothesis testing, the second of two general areas of

statistical inference. Hypothesis testing is a topic with which you as a student are

likely to have