Interval Estimation

Distribution of the Sample Mean

Wanted: More precise information on estimated parameters.

Our goal now is to gain more precise information on the value of an estimated parameter. What we have obtained so far are point estimates, but we do not yet know how close such an estimate is to the actual value of the parameter.

Needed: the distribution of the sample statistic.

As a first example, let us consider the sample mean:

13.1. Theorem. Let $X_1, \ldots, X_n$ be a random sample of size n ***from a normal distribution*** with mean μ and variance σ^2.

Then $\overline{X}$ is normally distributed with mean μ and variance σ^2/n.

13.2. Remark. Even if the sample is taken from a non-normal distribution, if n is “sufficiently large”, then the distribution of $\overline{X}$ will be close to normal due to the Central Limit Theorem 7.13.

Interval Estimation

13.3. Notation. We will often denote an interval of the form $[x-\varepsilon, x+\varepsilon]$ for $x \in \mathbb{R}$, $\varepsilon > 0$ by $x \pm \varepsilon$. In fact, we define

$$y = x \pm \varepsilon \qquad \text{to mean} \qquad y \in [x-\varepsilon, x+\varepsilon].$$

We would like to make statements such as "based on the results of a sample, we are 90% certain that the mean of a population lies in $\overline{X} \pm L$."

This is known as ***interval estimation*** and the resulting interval is called a ***confidence interval***.

Two-Sided Confidence Intervals

13.4. Definition. Let $0 \le \alpha \le 1$. A $100(1-\alpha)\%$ ***(two-sided) confidence interval for a parameter*** θ is an interval $[L_1, L_2]$ such that

$$P[L_1 \le \theta \le L_2] = 1 - \alpha. \tag{13.1}$$

13.5. Remark. The equation (13.1) does not determine L_1 and L_2 uniquely; we will nearly always require ***centered confidence intervals*** with

$$P[\theta < L_1] = P[\theta > L_2] = \alpha/2.$$

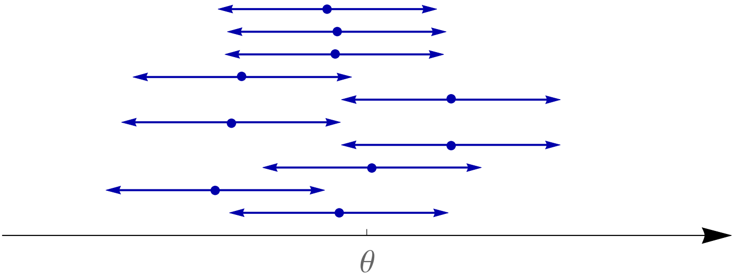
If the distribution of $\widehat{\theta}$ is symmetric about θ, then

$$L_1 = \widehat{\theta} - L, \qquad L_2 = \widehat{\theta} + L,$$

where L is a sample statistic and the interval is centered on $\hat{\theta}$, the point estimate for θ.

Random Intervals

13.6. Remark. It is important to note that in (13.1), the population parameter ***θ is not random***, but that ***L_1 and L_2 are random***. Hence, we may say that $[L_1, L_2]$ is a random interval.



Given θ, a random sample has a probability of $1-\alpha$ of yielding sample statistics L_1 and L_2 such that $\theta \in [L_1, L_2]$.

Interval Estimation for the Mean (Variance Known)

Suppose that we have a random sample of size n from a normal population with ***unknown mean*** μ and ***known variance*** σ^2.

A sample yields a point estimate $\overline{X}$ for μ. We want to find $L = L(\alpha)$ such that we can state with $100(1-\alpha)\%$ confidence that $\mu = \overline{X} \pm L$.

In particular, we would like to find a number L so that

$$P[\overline{X} - L \le \mu \le \overline{X} + L] = 1 - \alpha.$$

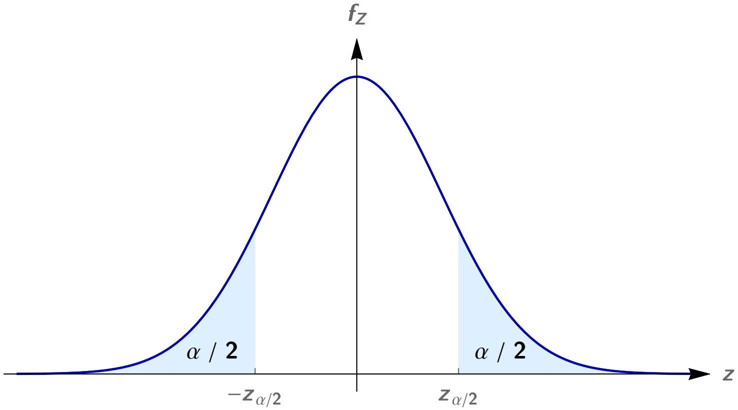
Note again that μ is not random, but rather a fixed but unknown parameter. However, the sample statistic $\overline{X}$ is random and so is L.

It is crucial that we know the distribution of $\overline{X}$.

The Point $z_{\alpha/2}$

Given $\alpha \in [0, 1]$ we define $z_{\alpha/2} \in [0, \infty)$ by

$$\alpha/2 = P[Z \geq z_{\alpha/2}] = \frac{1}{\sqrt{2\pi}} \int_{z_{\alpha/2}}^{\infty} e^{-x^2/2}\, dx. \tag{13.2}$$



Interval Estimation for the Mean (Variance Known)

Fix $\alpha \in [0,1]$. Then

$$1-\alpha = P[\overline{X} - L \le \mu \le \overline{X} + L] = P\left[\frac{\overline{X}-\mu-L}{\sigma/\sqrt{n}} \le 0 \le \frac{\overline{X}-\mu+L}{\sigma/\sqrt{n}}\right]$$

By Theorem 13.1 the sample mean is normally distributed with mean μ and variance σ^2/n. Thus,

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

follows a standard normal distribution, and so

$$\begin{aligned}
1-\alpha &= P\left[Z - \frac{L}{\sigma/\sqrt{n}} \le 0 \le Z + \frac{L}{\sigma/\sqrt{n}}\right] \\
&= P\left[-\frac{L}{\sigma/\sqrt{n}} \le Z \le \frac{L}{\sigma/\sqrt{n}}\right] \\
&= 2P\left[0 \le Z \le \frac{L}{\sigma/\sqrt{n}}\right] = 1 - 2P\left[\frac{L}{\sigma/\sqrt{n}} \le Z < \infty\right]
\end{aligned}$$

Confidence Interval for the Mean (Variance Known)

In this way we determine L as being the number such that

$$P\left[\frac{L}{\sigma/\sqrt{n}} \le Z < \infty\right] = \alpha/2.$$

This is equivalent to writing

$$\frac{L}{\sigma/\sqrt{n}} = z_{\alpha/2} \qquad \text{or} \qquad L = \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}.$$

We have proved the following result:

13.7. Theorem. Let $X_1, \dots, X_n$ be a random sample of size n from a normal distribution with mean μ and variance σ^2. A $100(1-\alpha)\%$ confidence interval on μ is given by

$$\overline{X} \pm \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}.$$

Confidence Interval for the Mean (Variance Known)

13.8. Example. An article in the ***Journal of Heat Transfer*** describes a method of measuring the thermal conductivity of Armco iron. Using a temperature of 100° F and a power input of 550 W, the following 10 measurements of thermal conductivity (in Btu /(hr ft ° F)) were obtained:

41.60	41.48	42.34	41.95	41.86
42.18	41.72	42.26	41.81	42.04

A point estimate of the mean thermal conductivity at 100° F and 550 W is the sample mean,

$$\overline{x} = 41.92\,\text{Btu}/(\text{hr ft}\ {}^\circ\text{F}).$$

Suppose we know that the standard deviation of the thermal conductivity under the given conditions is $\sigma = 0.10\,\text{Btu}/(\text{hr ft}\ {}^\circ\text{F})$. A 95% confidence interval ($\alpha = 0.05$) on the mean is then given by

$$\overline{x} \pm \frac{z_{0.025} \cdot \sigma}{\sqrt{n}} = 41.924 \pm \frac{1.96 \cdot 0.1}{\sqrt{10}} = [41.862, 41.986].$$

One-Sided Confidence Intervals

13.9. Definition. Let $0 \le \alpha \le 1$. A $100(1-\alpha)\%$ ***upper confidence bound*** for θ is a number L such that

$$P[\theta \le L] = 1 - \alpha.$$

A $100(1-\alpha)\%$ ***lower confidence bound*** for θ is a number L such that

$$P[L \le \theta] = 1 - \alpha.$$

The corresponding intervals are called ***one-sided confidence intervals***.

13.10. Theorem. Let $X_1, \dots, X_n$ be a random sample of size n from a normal distribution with mean μ and variance σ^2.

(i) A $100(1-\alpha)\%$ upper confidence bound on μ is given by $\overline{X} + \dfrac{z_\alpha \cdot \sigma}{\sqrt{n}}$.

(ii) A $100(1-\alpha)\%$ lower confidence bound on μ is given by $\overline{X} - \dfrac{z_\alpha \cdot \sigma}{\sqrt{n}}$.

Interval Estimation

Mathematica has built-in functionality for two-sided confidence intervals:

```
Needs["HypothesisTesting`"]

data := {41.60, 41.48, 42.34, 41.95, 41.86,
  42.18, 41.72, 42.26, 41.81, 42.04}

Mean[data]

41.924

MeanCI[data, KnownVariance → 0.01, ConfidenceLevel → 0.95]

{41.862, 41.986}
```

The value for $z_{\alpha/2}$ may be found by inverting the cumulative distribution function. For instance, for $\alpha = 0.05$,

```
InverseCDF[NormalDistribution[0, 1], 0.975]

1.95996
```

This is useful for finding one-sided confidence intervals.

Joint Sampling of Mean and Variance

Our interest in the chi-squared distribution is not merely abstract, for understanding the sum of squares of normally distributed random variables; in fact, the main application lies in analyzing the distribution of the sample variance. In the previous chapter, we were able to analyze the sample mean, and also its distribution, under the assumption of ***known variance***. If the variance

$$\sigma^2 = \mathrm{E}[(X-\mu)^2]$$

is unknown, we must start all over again, and first learn more about the sample variance

$$S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2.$$

The problem essentially is that we are using the random sample $X_1, \dots, X_n$ to obtain $\overline{X}$ and S^2 at the same time, i.e., we actually need to obtain the ***joint distribution*** of $\overline{X}$ and S^2.

Joint Distribution of Sample Mean and Variance

The following theorem and the chi-squared distribution were discovered by Helmert in 1876 in the context of statistics of geodesical measurements. It was published in German textbooks.

Friedrich Robert Helmert (1843-1917).
File:F-R Helmert 1.jpg. (2016, January 27). Wikimedia Commons, the free media repository.

However, his results were unknown to English statisticians and the chi-squared distribution was rediscovered by Pearson in 1900. Fisher and Grosset (see below) then found its application to statistics.

13.11. Theorem. Let $X_1, \dots, X_n$, $n \geq 2$, be a random sample of size n from a normal distribution with mean μ and variance σ^2. Then

(i) The sample mean $\overline{X}$ is independent of the sample variance S^2,

(ii) $\overline{X}$ is normally distributed with mean μ and variance σ^2/n,

(iii) $(n-1)S^2/\sigma^2$ is chi-squared distributed with $n-1$ degrees of freedom.

The Helmert Transformation

The ***Helmert transformation*** is a very special kind of ***linear, orthogonal map*** from a set of $n \geq 2$ i.i.d. normal random variables $X_1, \dots, X_n$ to a new set of random variables $Y_1, \dots, Y_n$.

A sample of size n taken from a normal population X with mean μ and variance σ^2 is transformed as follows:

$$\begin{aligned}
Y_1 &= \frac{1}{\sqrt{n}}(X_1 + \dots + X_n) \\
Y_2 &= \frac{1}{\sqrt{2}}(X_1 - X_2) \\
Y_3 &= \frac{1}{\sqrt{6}}(X_1 + X_2 - 2X_3) \\
&\vdots \\
Y_n &= \frac{1}{\sqrt{n(n-1)}}(X_1 + X_2 + \dots + X_{n-1} - (n-1)X_n)
\end{aligned} \tag{13.3}$$

The Helmert Transformation

In matrix notation,

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & -\frac{n-1}{\sqrt{n(n-1)}} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_n \end{pmatrix}$$

or $\mathbf{Y} = A\mathbf{X}$ for short. It is easy to see that the rows of the matrix A are orthonormal. Thus, A is an orthogonal matrix, $A^{-1} = A^T$. This immediately implies $|\det A| = 1$, since

$$\det A = \det A^T = \det A^{-1} = \frac{1}{\det A} \quad \Rightarrow \quad (\det A)^2 = 1$$

The Helmert Transformation

Incidentally, the orthogonality of A also implies that if $\mathbf{y} = A\mathbf{x}$, then

$$\sum_{i=1}^{n} y_i^2 = \langle \mathbf{y}, \mathbf{y} \rangle = \langle A\mathbf{x}, A\mathbf{x} \rangle = \langle A^T A\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2. \qquad (13.4)$$

We have assumed that the random variables $X_1, \dots, X_n$ are i.i.d., so their joint distribution function is given by the product of the individual normal distributions,

$$\begin{aligned} f_{X_1 \cdots X_n}(x_1, \dots, x_n) &= \prod_{i=1}^{n} (2\pi)^{-1/2} \sigma^{-1} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2} \\ &= (2\pi)^{-n/2} \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i^2 - 2\mu x_i + \mu^2)} \\ &= (2\pi)^{-n/2} \sigma^{-n} e^{-\frac{1}{2\sigma^2} \left(\sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2 \right)} \end{aligned}$$

The Helmert Transformation

The Helmert transformation is linear, so its derivative (Jacobian) DA is simply A. Using (13.4), $|\det A^{-1}| = 1$ and Theorem 10.1 on the transformation of joint random variables, we obtain

$$\begin{aligned}
& f_{Y_1 \cdots Y_n}(y_1, \dots, y_n) \\
&= f_{Y_1 \cdots Y_n}(\mathbf{y}) = f_{X_1 \cdots X_n}(\mathbf{x})_{\mathbf{x}=A^{-1}\mathbf{y}} \cdot \underbrace{|\det DA^{-1}(\mathbf{y})|}_{=1} \\
&= (2\pi)^{-n/2}\sigma^{-n} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n y_i^2 - 2\mu\sqrt{n}y_1 + n\mu^2\right)} \\
&= (2\pi)^{-n/2}\sigma^{-n} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=2}^n y_i^2 + (y_1 - \sqrt{n}\mu)^2\right)} \\
&= (2\pi)^{-1/2}\sigma^{-1} e^{-\frac{1}{2\sigma^2}(y_1-\sqrt{n}\mu)^2} \prod_{i=2}^n (2\pi)^{-1/2}\sigma^{-1} e^{-\frac{1}{2\sigma^2}y_i^2}
\end{aligned}$$

The Helmert Transformation

We see that the marginal densities are given by

$$f_{Y_1}(y_1) = (2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}(y_1-\sqrt{n}\mu)^2},$$

$$f_{Y_i}(y_i) = (2\pi)^{-1/2}\sigma^{-1}e^{-\frac{1}{2\sigma^2}y_i^2}, \qquad i = 2, \dots, n$$

and the joint density is the product of the marginal densities,

$$f_{Y_1\cdots Y_n}(y_1, \dots, y_n) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) \dots f_{Y_n}(y_n).$$

In particular, the random variables $Y_1, \dots, Y_n$ are ***independent*** and ***normally distributed***.

The random variable Y_1 is normally distributed with mean $\sqrt{n}\mu$ and variance σ^2, while $Y_2, \dots, Y_n$ have mean 0 and variance σ^2.

The Helmert Transformation

Proof of Theorem 13.11.

Using the Helmert transformation, we may write

$$\overline{X} = \frac{1}{\sqrt{n}} Y_1.$$

Furthermore,

$$\begin{aligned}(n-1)S^2 &= \sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 = \sum_{i=1}^{n} Y_i^2 - Y_1^2 \\ &= \sum_{i=2}^{n} Y_i^2.\end{aligned}$$

Since the Y_i are all independent, it follows that $\overline{X}$ is independent of S^2, so we have proven the first sassertion of the theorem.

The Helmert Transformation

Proof of Theorem 13.11 (continued).

Since $\overline{X} = \frac{1}{\sqrt{n}} Y_1$ and $f_{Y_1}(y_1) = (2\pi)^{-1/2}\sigma^{-1} e^{-\frac{(y_1 - \sqrt{n}\mu)^2}{2\sigma^2}}$, it follows from Theorem 7.5 that

$$f_{\overline{X}}(x) = (2\pi)^{-1/2}\sigma^{-1} e^{-\frac{(\sqrt{n}x - \sqrt{n}\mu)^2}{2\sigma^2}} \sqrt{n}$$

so $\overline{X}$ is normally distributed with mean μ and variance σ^2/n.

Now

$$(n-1)S^2/\sigma^2 = \frac{1}{\sigma^2}\sum_{i=2}^{n} Y_i^2 = \sum_{i=2}^{n}\left(\frac{Y_i}{\sigma}\right)^2$$

is the sum of $n-1$ squares of standard normal distributions Y_i/σ, so it follows a chi-squared distribution with $n-1$ degrees of freedom.

This completes the proof. $\square$

Independence of Sample Mean and Sample Variance

13.12. Remark. Theorem 13.11 essentially uses the fact that the i.i.d. variables X_k, $k = 1, \dots, n$, are normally distributed. In fact, the converse result is true also:

> *Let $X_1, \dots, X_n$, $n \geq 2$, be i.i.d. random variables. Then if $\overline{X}$ and S^2 are independent, the X_k, $k = 1, \dots, n$ follow a normal distribution.*

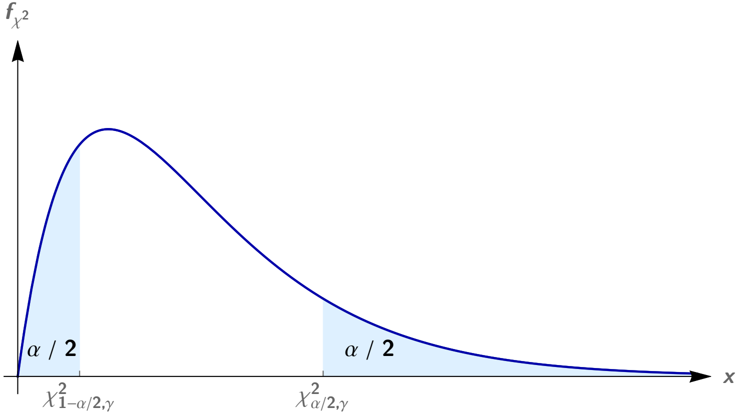
This means that the independence of $\overline{X}$ and S^2 is a ***characteristic property*** of the normal distribution. Furthermore, if in a given situation we assume that $\overline{X}$ and S^2 are independently distributed we are essentially assuming that the population is normally distributed.

We can use Theorem 13.11 to find a confidence interval for the variance based on the sample variance S^2.

The Points $\chi^2_{1-\alpha/2,\gamma}$ and $\chi^2_{\alpha/2,\gamma}$

Given $\alpha \in [0,1]$ and $\gamma > 0$ we define $\chi^2_{1-\alpha/2,\gamma}, \chi^2_{\alpha/2,\gamma} \in [0,\infty)$ by

$$\int_0^{\chi^2_{1-\alpha/2,n}} f_{\chi^2_n}(x)\, dx = \int_{\chi^2_{\alpha/2,n}}^{\infty} f_{\chi^2_n}(x)\, dx = \alpha/2,$$



Interval Estimation of Variability

From Theorem 13.11 we know that given a sample of size n from a normal population, $(n-1)S^2/\sigma^2$ follows a chi-squared distribution with $n-1$ degrees of freedom. Thus

$$\begin{aligned}1-\alpha &= P\left[\chi^2_{1-\alpha/2,n-1} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{\alpha/2,n-1}\right] \\ &= P\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right]\end{aligned}$$

This gives us the following result:

13.13. Theorem. Let $X_1, \ldots, X_n$, $n \ge 2$, be a random sample of size n from a normal distribution with mean μ and variance σ^2. A $100(1-\alpha)\%$ confidence interval on σ^2 is given by

$$\left[(n-1)S^2/\chi^2_{\alpha/2,n-1},\ (n-1)S^2/\chi^2_{1-\alpha/2,n-1}\right].$$

Interval Estimation of Variability

Often, we are only interested in finding an upper or lower bound for the variance.

13.14. Theorem. Let $X_1, \ldots, X_n$, $n \geq 2$, be a random sample of size n from a normal distribution with mean μ and variance σ^2. Then with $100(1-\alpha)\%$ confidence,

$$\sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha,n-1}}$$

and $[0, \frac{(n-1)S^2}{\chi^2_{1-\alpha,n-1}}]$ is a $100(1-\alpha)\%$ ***upper confidence interval*** for σ^2.

Similarly, with $100(1-\alpha)\%$ confidence

$$\frac{(n-1)S^2}{\chi^2_{\alpha,n-1}} \leq \sigma^2.$$

and $[\frac{(n-1)S^2}{\chi^2_{\alpha,n-1}}, \infty)$ is a $100(1-\alpha)\%$ ***lower confidence interval*** for σ^2.

Interval Estimation of Variability

13.15. Example. A manufacturer of soft drink beverages is interested in the uniformity of the machine used to fill cans. Specifically, it is desirable that the standard deviation σ of the filling process be less than 0.2 fluid ounces; otherwise there will be a higher than allowable percentage of cans that are underfilled. We will assume that fill volume is approximately normally distributed. A random sample of 20 cans results in a sample variance of $s^2 = 0.0225\,(\text{fluid ounces})^2$. A 95% upper-confidence interval is given by

$$\sigma^2 \le \frac{(n-1)S^2}{\chi^2_{0.95,n-1}} = \frac{19 \cdot 0.0225\,(\text{fluid ounces})^2}{10.117} = 0.0423\,(\text{fluid ounces})^2$$

This corresponds to $\sigma \le 0.21$ fluid ounces with 95% confidence. This is not sufficient to support the hypothesis that $\sigma \le 0.20$ fluid ounces so further investigation is necessary.

Interval Estimation of Variability

Mathematica has built-in functionality for two-sided confidence intervals for the variance:

```
data := {41.60, 41.48, 42.34, 41.95, 41.86,
  42.18, 41.72, 42.26, 41.81, 42.04}

VarianceCI[data, ConfidenceLevel -> .95]

{0.0381879, 0.269013}
```

However, one-sided intervals need to be calculated by hand, using, for example,

```
InverseCDF[ChiSquareDistribution[19], 0.05]

10.117
```

Interval Estimation for the Mean (Variance unknown)

Recall that we have derived a formula for the confidence interval of the mean of a normal distribution using the random variable

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

which was found to be normally distributed. The Central Limit Theorem allowed us to extend this result (approximately) even to non-normal distributions, but one central difficulty remained: σ must be known!

Our main goal is to derive a general formula for a confidence interval on the mean when he value of σ is not known and must be estimated.

The difficulty lies in the fact that the distribution of

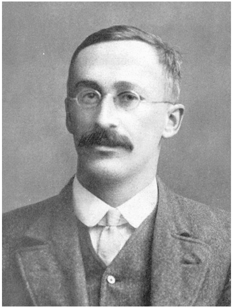
$$\frac{\overline{X} - \mu}{S/\sqrt{n}}$$

is not known.

The Student T-distribution

Gosset was a statistician and brewer who worked fur the Guinness company in Dublin. He was interested in developing new varieties of barley and became a pioneer of many important statistical methods.

To prevent the leaking of trade secrets, Guinness prohibited their staff from publishing findings that mentioned beer, "Guinness" or the author's name. Therefore, Gosset published his results under the pseudonym "A. Student."



William Sealy Gosset (1876-1937) in 1908.
File:William Sealy Gosset.jpg. (2017, April 26). Wikimedia Commons, the free media repository.

13.16. Definition. Let Z be a standard normal variable and let χ_γ^2 be an ***independent*** chi-squared random variable with γ degrees of freedom. The random variable

$$T_\gamma = \frac{Z}{\sqrt{\chi_\gamma^2/\gamma}}$$

is said to follow a T-distribution with γ degrees of freedom.

Density of the T-distribution

13.17. Theorem. The density of a T distribution with γ degrees of freedom is given by

$$f_{T_\gamma}(t) = \frac{\Gamma((\gamma+1)/2)}{\Gamma(\gamma/2)\sqrt{\pi\gamma}} \left(1 + \frac{t^2}{\gamma}\right)^{-\frac{\gamma+1}{2}} .$$

Proof.

The distribution of χ_γ was found in (10.1) to be

$$f_{\chi_\gamma}(y) = \begin{cases} \frac{2}{2^{\gamma/2}\Gamma(\gamma/2)} y^{\gamma-1} e^{-\gamma^2/2} & y \geq 0, \\ 0 & y < 0. \end{cases}$$

It follows from Theorem 7.5 that $\sqrt{\chi_\gamma^2/\gamma} = \chi_\gamma/\sqrt{\gamma}$ has distribution

$$f_{\chi_\gamma/\sqrt{\gamma}}(y) = \begin{cases} \frac{2\sqrt{\gamma}}{2^{\gamma/2}\Gamma(\gamma/2)} (\sqrt{\gamma}y)^{\gamma-1} e^{-\gamma y^2/2} & y \geq 0, \\ 0 & y < 0. \end{cases}$$

Density of the T-distribution

Proof (continued).
The density of the standard normal random variable Z is given by

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

By Theorem 10.2, the density f_T of the quotient $T = X/Y$ of two independent random variables X and Y is given by

$$f_T(t) = \int_{-\infty}^{\infty} f_X(ty) f_Y(y) \cdot |y| \, dy.$$

For $T_\gamma = Z/(\chi_\gamma/\sqrt{\gamma})$ it follows that

$$f_{T_\gamma}(t) = \frac{1}{\sqrt{2\pi}} \frac{2\sqrt{\gamma}}{2^{\gamma/2}\Gamma(\gamma/2)} \int_0^{\infty} e^{-(t^2+\gamma)y^2/2} (\sqrt{\gamma} y)^{\gamma-1} y \, dy.$$

Density of the T-distribution

Proof (continued).
Substituting $y = \sqrt{2z/(t^2+\gamma)}$, $z = (t^2+\gamma)y^2/2$, $dz = (t^2+\gamma)y\,dy$, we obtain

$$\begin{aligned} f_{T_\gamma}(t) &= \frac{1}{\sqrt{2\pi}} \frac{2\sqrt{\gamma}}{2^{\gamma/2}\Gamma(\gamma/2)}(t^2+\gamma)^{-1} \int_0^\infty e^{-z}\left(\frac{2z\gamma}{t^2+\gamma}\right)^{\frac{\gamma-1}{2}}\,dz \\ &= \frac{1}{\sqrt{\gamma\pi}} \frac{1}{\Gamma(\gamma/2)} \left(\frac{\gamma}{t^2+\gamma}\right)^{\frac{\gamma+1}{2}} \int_0^\infty e^{-z} z^{\frac{\gamma+1}{2}-1}\,dz \\ &= \frac{1}{\sqrt{\gamma\pi}} \frac{\Gamma((\gamma+1)/2)}{\Gamma(\gamma/2)} \left(1+\frac{t^2}{\gamma}\right)^{-\frac{\gamma+1}{2}}. \end{aligned}$$

$\square$

T-distribution of the Sample Mean

13.18. Theorem. Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2. The random variable

$$T_{n-1} = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

follows a T distribution with $n-1$ degrees of freedom.

Proof.
We know that $(\overline{X} - \mu)/(\sigma/\sqrt{n})$ is standard normal and $(n-1)S^2/\sigma^2$ is a chi-squared random variable with $n-1$ degrees of freedom. Therefore,

$$\frac{Z}{\sqrt{\chi^2_\gamma/\gamma}} = \frac{(\overline{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{((n-1)S^2/\sigma^2)/(n-1)}} = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

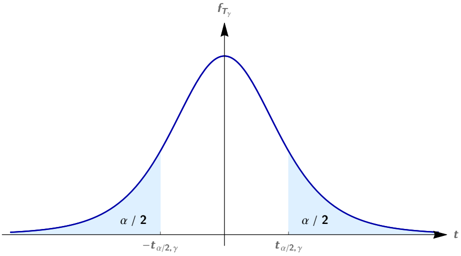
follows a T distribution with $n-1$ degrees of freedom. $\square$

Confidence Interval for the Mean (Variance Unknown)

Let $0 < \alpha \le 1$ and $\gamma > 0$. We define $t_{\alpha/2,\gamma} \ge 0$ by

$$\int_{t_{\alpha/2,\gamma}}^{\infty} f_{T_\gamma}(t)\, dt = \alpha/2, \quad (13.5)$$

where f_{T_γ} is the density of the T-distribution with n degrees of freedom.



13.19. Theorem. Let $X_1, \dots, X_n$ be a random sample of size n ***from a normal distribution*** with mean μ and variance σ^2. Then a $100(1-\alpha)\%$ confidence interval on μ is given by

$$\overline{X} \pm t_{\alpha/2,n-1} S/\sqrt{n}$$

Confidence Interval for the Mean (Variance Unknown)

13.20. Example. An article in the ***Journal of Testing and Evaluation*** presents the following 20 measurements on residual flame time (in seconds) of treated specimens of children's nightwear:

9.85	9.93	9.75	9.77	9.67	9.87	9.67	9.94	9.85	9.75
9.83	9.92	9.74	9.99	9.88	9.95	9.95	9.93	9.92	9.89

We wish to find a 95% confidence interval on the mean residual flame time. The sample mean and standard deviation are

$$\overline{x} = 9.8475, \qquad s = 0.0954$$

We refer to the table for the T distribution with $20 - 1 = 19$ degrees of freedom and and $\alpha/2 = 0.025$ to obtain $t_{0.025,19} = 2.093$. Hence

$$\mu = (9.8475 \pm 0.0446)\,\text{sec}, \qquad \text{i.e.,} \qquad 9.8029 \leq \mu \leq 9.8921$$

with 95% probability.