Pruebas de hipótesis e Intervalos de confianza

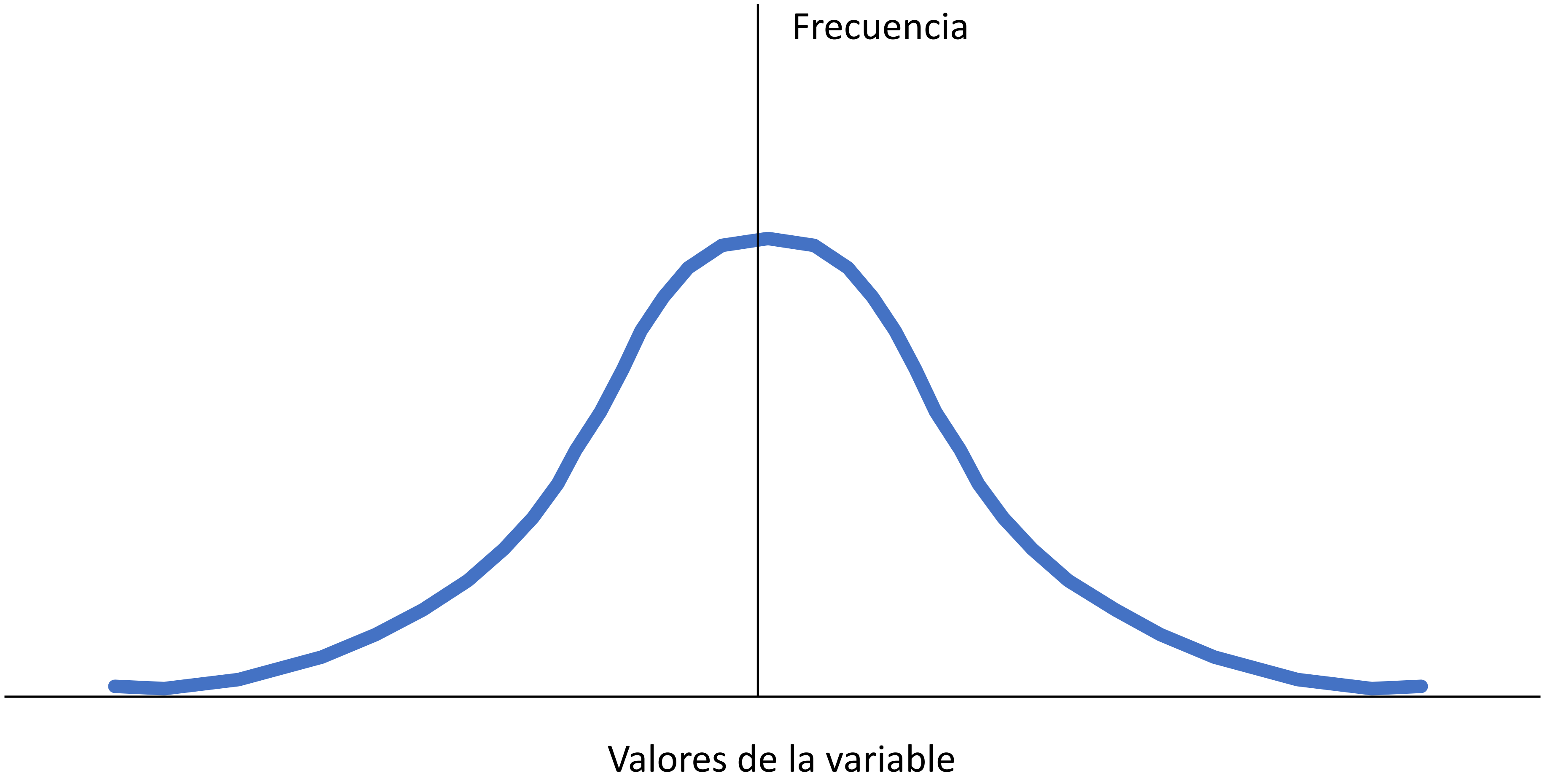
El carácter dual de las muestras

Sea $x_1, \ldots, x_n$ una muestra aleatoria de una variable X.

La muestra puede considerse de dos formas

- Es un conjunto de variables con la misma distribución de X cuando la muestra fue seleccionada pero todavía no se realizaron las mediciones
- Es un conjunto de números correspondientes a las mediciones hechas sobre los individuos de la muestra

El carácter dual de las muestras



La muestra como un conjunto de variables normales

Sea $x_1, \dots, x_n$ una muestra aleatoria de una variable X.

- Si X tiene distribución $N(\mu, \sigma)$, entonces todas las $x_i \sim N(\mu, \sigma)$
- El estimador óptimo de μ es el promedio

$$\hat{\mu} = \bar{X}$$

- El estimador óptimo de σ es el desvío standard

$$\hat{\sigma} = S = \sqrt{\frac{\sum (x_i - \bar{X})^2}{n - 1}}$$

La muestra como un conjunto de variables normales

Sea $x_1, \ldots, x_n$ una muestra aleatoria de una variable X.

- Si X tiene distribución $N(\mu, \sigma)$, entonces la distribución del promedio es

$$\bar{X} \sim N\left(\mu, {}^{\sigma}/_{\sqrt{n}}\right)$$

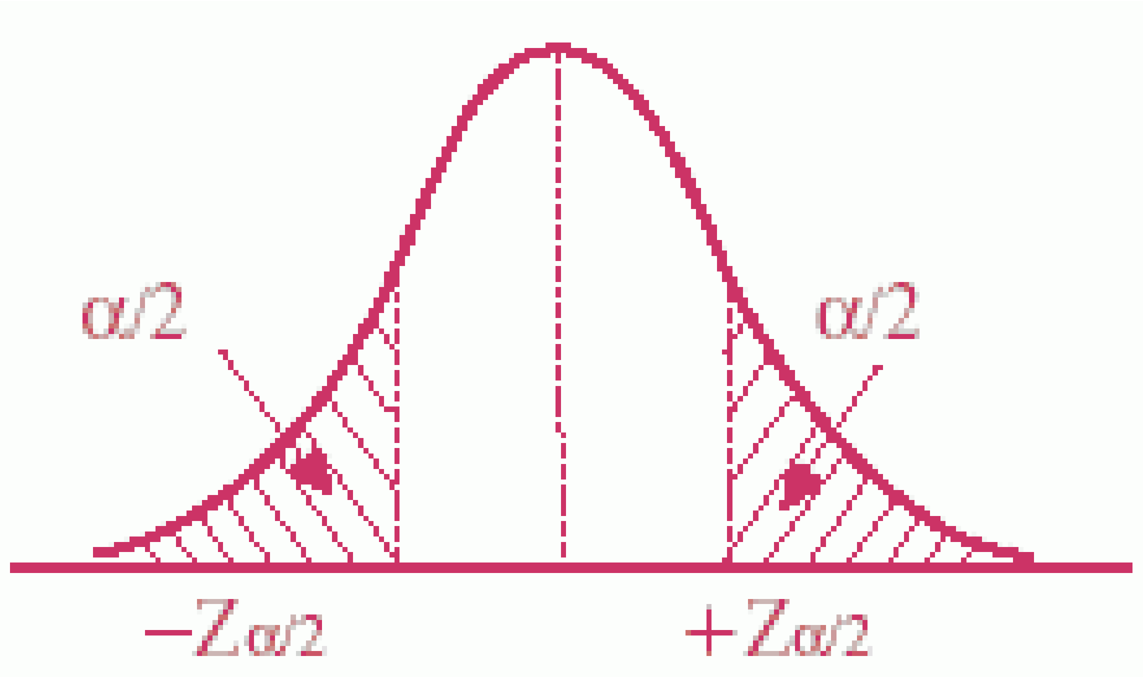
$$\frac{\bar{X} - \mu}{{}^{\sigma}/_{\sqrt{n}}} \sim N(0,1)$$

Intervalo de confianza

- Es una forma alternativa de estimar un parámetro
- Consiste en un rango de valores el cual contendrá al valor poblacional con una alta probabilidad
- Por ejemplo, el intervalo de confianza para la media de una población se extiende a ambos lados de la media muestral un número fijo de desvíos stándard.

Intervalo de confianza para la media de una normal

$$\bar{X} \sim N\left(\mu, \sigma/\sqrt{n}\right)$$



$$P\left(-z_{\alpha/2} \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(\bar{X} - z_{\alpha/2}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

Intervalo de confianza del 95%

- Es cuando $\alpha = 0.05$
- En este caso, es de esperar que el IC contenga al parámetro poblacional el 95% de las veces

- Límite inferior: $\bar{X} - 1.96\frac{s}{\sqrt{n}}$
- Límite superior $\bar{X} + 1.96\frac{s}{\sqrt{n}}$

- Dado que $z_{0.05/_2} = z_{0.025} = 1.96$

Intervalo de confianza para un estimador

- En general, para casi cualquier parámetro se puede armar un IC con un estimador y su desvío standard (SD):

$$estim - 2SD; estim + 2SD$$

- Este resultado se basa en el teorema central del límite

Prueba de Hipótesis

Prueba de hipótesis

Es habitual que en una investigación se pretenda determinar la veracidad de una afirmación en la población de referencia

- Ejemplos de afirmaciones cuya veracidad interesa evaluar
- El colesterol sérico es más alto en los hombres que en las mujeres?
- El tratamiento contra el insomnio es efectivo?

Test de hipótesis

- En cada caso, lo que se pretende es evaluar la veracidad de estas afirmaciones en la muestra disponible y que las conclusiones obtenidas sean válidas para la población de pertenencia
- Para que esto sea así, la muestra debe ser representativa de la población
- Y deberán aplicarse métodos que permitan decidir sobre la veracidad de una afirmación con un cierto nivel de error

Test de hipótesis

- Es el instrumento estadístico que permite evaluar a partir de los datos de una muestra, la veracidad de afirmaciones hechas sobre una población
- Problema: sacando conclusiones con informacion incompleta (tomando decisiones) se pueden cometer errores
- Objetivo: diseñar una estrategia que mantenga dichos errores al mínimo.

Componentes de un Test de hipótesis

- En toda investigación, la hipótesis a evaluar tiene una hipótesis complementaria. Sólo una de ellas puede ser verdadera en la población
- En la teoría de test de hipótesis tenemos
  - Hipótesis nula H_0
  - Hipótesis alternativa, complementaria de la anterior, H_1

Test de hipótesis

Dos posibles errores pueden cometerse cuando se utiliza esta metodología de decisión basada en datos muestrales

- Error de tipo I: Rechazar Ho cuando es verdadera
- Error de tipo II: Aceptar Ho cuando es falsa

Tipos de error

Test	Población	
	Ho Verd.	Ho Falsa
No rechaza Ho	OK	Error tipo II Prob error = β
Rechaza Ho	Error tipo I Prob error = α	OK

Componentes de una prueba de hipótesis

- **Hipótesis nula** $\mathbf{H_0}$
- **Hipótesis alternativa**, complementaria de la anterior, $\mathbf{H_1}$

- **Error de tipo I**: Rechazar Ho cuando es verdadera
- **Error de tipo II**: Aceptar Ho cuando es falsa

Componentes de una prueba de hipótesis

- **Nivel de significación**
  - Probabilidad de error de tipo I = α
  - α se fija a priori (0.05, 0.01 son los valores habituales)
- **Probabilidad de error de tipo II = β**
  - β depende del tamaño de la muestra considerada
- **Potencia del test = 1- β**
  - Se puede calcular el tamaño de la muestra para obtener una potencia determinada

Prueba de hipótesis

Estadístico de Prueba de un test

- Se calcula con los datos de la muestra
- Tiene distribución conocida cuando Ho es verdadera
- Esto es necesario para poder diseñar la prueba de hipótesis de manera que la probabilidad de error de tipo I sea la deseada

Prueba de hipótesis

Región de rechazo de un test

Es el conjunto de valores del estadístico de prueba para los cuales se rechaza H_0

Región de aceptación de un test

- Es el conjunto de valores del estadístico de prueba para los cuales no se rechaza H_0

Prueba de hipótesis

REGLA DE DECISION

1. Se fija un valor de α. Máximo error de decisión tolerado
2. Se calcula la probabilidad de equivocarse al rechazar Ho, con los datos de la muestra (Valor-p ó p-value) (otros nombres: sig., signif.)

- **RECHAZO Ho** si el **valor-p** es menor o igual a α:
- **NO RECHAZO Ho** si el **valor-p** no es menor a α

Problema

- El bajo peso al nacer, definido por un peso al nacer inferior a 2500 gr., ha sido una preocupación de los médicos durante años debido a que tanto las tasas de mortalidad como la de nacimientos defectuosos son muy altas para los niños con este problema. El comportamiento de la mujer durante el embarazo (incluyendo la dieta, los hábitos tabáquicos y los cuidados prenatales) pueden alterar las chances de un parto de un niño con bajo peso.
- Los datos que se presentan en este ejercicio corresponden a 189 nacimientos de los cuales 59 han resultado en niños con bajo peso. El objetivo de este ejercicio es determinar cuáles de las variables presentes en la tabla de datos que se adjunta son factores de riesgo de bajo peso al nacer.
- Para iniciar los estudios de esta población de madres, se desea saber si la edad promedio de las madres es de 23 años

Situación

- Tenemos una magnitud medida en los mismos individuos y se necesita comparar el comportamiento de dicha medición contra un valor fijo, dado por la bibliografía
- Este es un problema de comparación de media poblacional contra un valor fijo

Ejemplo: Edad de la madre

Determinar si hay evidencia suficiente para afirmar que el promedio de edad es de 23 años

- H0: el promedio de edad es de 23 años
- H1: el promedio de edad no es de 23 años

Prueba de hipótesis para la media

$X_1, \ldots\ldots, X_n$ Muestra de una **población normal** con media μ y desvío σ

H_0 : **μ >= μo** H_1 : **μ < μo**

H_0 : **μ <= μo** H_1 : **μ > μo**

H_0 : **μ = μo** H_1 : **μ ≠ μo**

Test para la media de la normal

Estadístico de prueba

Sigma conocido

$$E = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

Sigma desconocido

$$E = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

Habitualmente se utiliza este último

Ejemplo: Edad de la madre

Determinar si hay evidencia suficiente para afirmar que el promedio de edad es de 23 años

- H0: el promedio de edad es de 23 años
- H1: el promedio de edad no es de 23 años

Ejemplo: Edad de la madre

- H0: el promedio de edad es de 23 años
- H1: el promedio de edad no es de 23 años

```
t.test(LOWBWT$AGE, mu=23, alternative = c("two.sided"))

        One Sample t-test

data:  LOWBWT$AGE
t = 0.61775, df = 188, p-value = 0.5375

alternative hypothesis: true mean is not equal to 23

95 percent confidence interval:
 22.47779 23.99840
sample estimates:
mean of x
  23.2381
```

```
describe(LOWBWT$AGE, na.rm=TRUE,
skew = F, ranges = F, trim=.1, quant = NULL, IQR = F)
   vars   n  mean  sd   se
X1    1 189 23.24 5.3 0.39
```

La función t.test() en R

```
t.test(x, y = NULL, alternative = c("two.sided", "less",
"greater"), mu = 0, paired = FALSE, var.equal = FALSE,
conf.level = 0.95, …)
```

Comparación de muestras independientes

Problema

- El bajo peso al nacer, definido por un peso al nacer inferior a 2500 gr., ha sido una preocupación de los médicos durante años debido a que tanto las tasas de mortalidad como la de nacimientos defectuosos son muy altas para los niños con este problema. El comportamiento de la mujer durante el embarazo (incluyendo la dieta, los hábitos tabáquicos y los cuidados prenatales) pueden alterar las chances de un parto de un niño con bajo peso.
- Los datos que se presentan en este ejercicio corresponden a 189 nacimientos de los cuales 59 han resultado en niños con bajo peso. El objetivo de este ejercicio es determinar cuáles de las variables presentes en la tabla de datos que se adjunta son factores de riesgo de bajo peso al nacer.
- Para iniciar los estudios de esta población de madres, se desea saber si la edad promedio de las madres que tuvieron nacimientos con bajo peso es diferente de la de aquellas que tuvieron parto normal

Situación

- Tenemos una magnitud (la edad) medida en dos muestras y se necesita compararlas para determinar si presentan evidencia de diferencias entre las poblaciones a las que pertenecen
- Este es un problema de comparación de medias de dos poblaciones independientes

Test de t para comparación de dos medias

$X_{11}, \ldots\ldots, X_{1n}$ Muestra de una población normal con media μ_1 y desvío σ_1

$X_{21}, \ldots\ldots, X_{2m}$ Muestra de una población normal con media μ_2 y desvío σ_2

$$H_a: \mu_1 - \mu_2 \neq \Delta_0$$

$$H_0: \mu_1 - \mu_2 = \Delta_0 \qquad H_a: \mu_1 - \mu_2 > \Delta_0$$

$$H_a: \mu_1 - \mu_2 < \Delta_0$$

Ho y Ha son eventos *exhaustivos y mutuamente excluyentes*.

Una de las dos afirmaciones *debe* ser verdadera.

Test de t para comparación de dos medias

Estadísticos de prueba

Sigmas desconocidos
e iguales

$$E = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$$

Sigmas desconocidos
y no iguales

$$E = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}} \sim t_{gl}$$

Test de t para comparación de dos medias

Supuestos de la prueba t para dos muestras independientes

- **Independencia**: Las observaciones tienen que ser independientes las unas de las otras. Para ello, el muestreo debe ser aleatorio
- **Normalidad**: Las poblaciones que se comparan tienen que seguir una distribución normal. En caso de cierta asimetría, los t-test son considerablemente robustos si el tamaño de las muestras es mayor o igual a 30.
- **Igualdad de varianza (homocedasticidad):** la varianza de las poblaciones comparadas debe de ser igual. En caso de no cumplirse esta condición se puede emplear un Welch Two Sample t-test

Es importante evaluar hasta que punto, el no cumplimiento de una o varias de sus condiciones, puede afectar al resultado del t-test. Si su utilización queda descartada, se puede recurrir a otros test estadísticos.

Ejemplo: Edad de la madre

Determinar si hay evidencia suficiente para afirmar que hay diferencias en el promedio de edad de las madres

- H0: no hay diferencia en el promedio de edad entre las madres de hijos con bajo peso y peso normal
- H1: hay diferencia en el promedio de edad entre las madres de hijos con bajo peso y peso normal

Ejemplo: Edad de la madre

- H0: no hay diferencia en el promedio de edad entre las madres de hijos con bajo peso y peso normal
- H1: hay diferencia en el promedio de edad entre las madres de hijos con bajo peso y peso normal

```
> t.test(LOWBWT$AGE~LOWBWT$LOW)

        Welch Two Sample t-test

data:  LOWBWT$AGE by LOWBWT$LOW
t = 1.7737, df = 136.94, p-value = 0.07834
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1558349  2.8687423
sample estimates:
mean in group 0 mean in group 1
       23.66154        22.30508
```

```
describeBy(LOWBWT$AGE,LOWBWT$LOW, na.rm=TRUE,
 skew = F, ranges = F, trim=.1, quant = NULL, IQR = F)

 Descriptive statistics by group
group: 0
   vars   n  mean   sd   se
X1    1 130 23.66 5.58 0.49
-----------------------------------------------------
group: 1
   vars  n  mean   sd   se
X1    1 59 22.31 4.51 0.59
```