

BIOESTADÍSTICA

Yasna Orellana Zapata

Profesora Asistente
Unidad de Nutrición Pública



Temas a tratar

- Teorema del límite central
- Estimador de la desviación estándar σ
- Estadístico T y Distribución T-Student
- Intervalo de Confianza para una media



Teorema del Límite Central

Si seleccionamos una muestra de tamaño n desde una población y nos interesa estudiar una variable cuantitativa, con distribución cualquiera F , que suponemos tiene una media poblacional μ , y una desviación estándar poblacional σ , (finitos) entonces:

El promedio muestral tiene distribución aproximadamente Normal con media μ y desviación estándar σ/\sqrt{n} .

Esta aproximación es mejor cuando $n > 30$.

En el caso en que σ es desconocido, es necesario estimarlo

Un estimador (insesgado) de la desviación estándar, σ es:

$$S = \sqrt{\left(\frac{1}{n-1} \times \sum_{i=1}^n (x_i - \bar{X})^2 \right)}$$

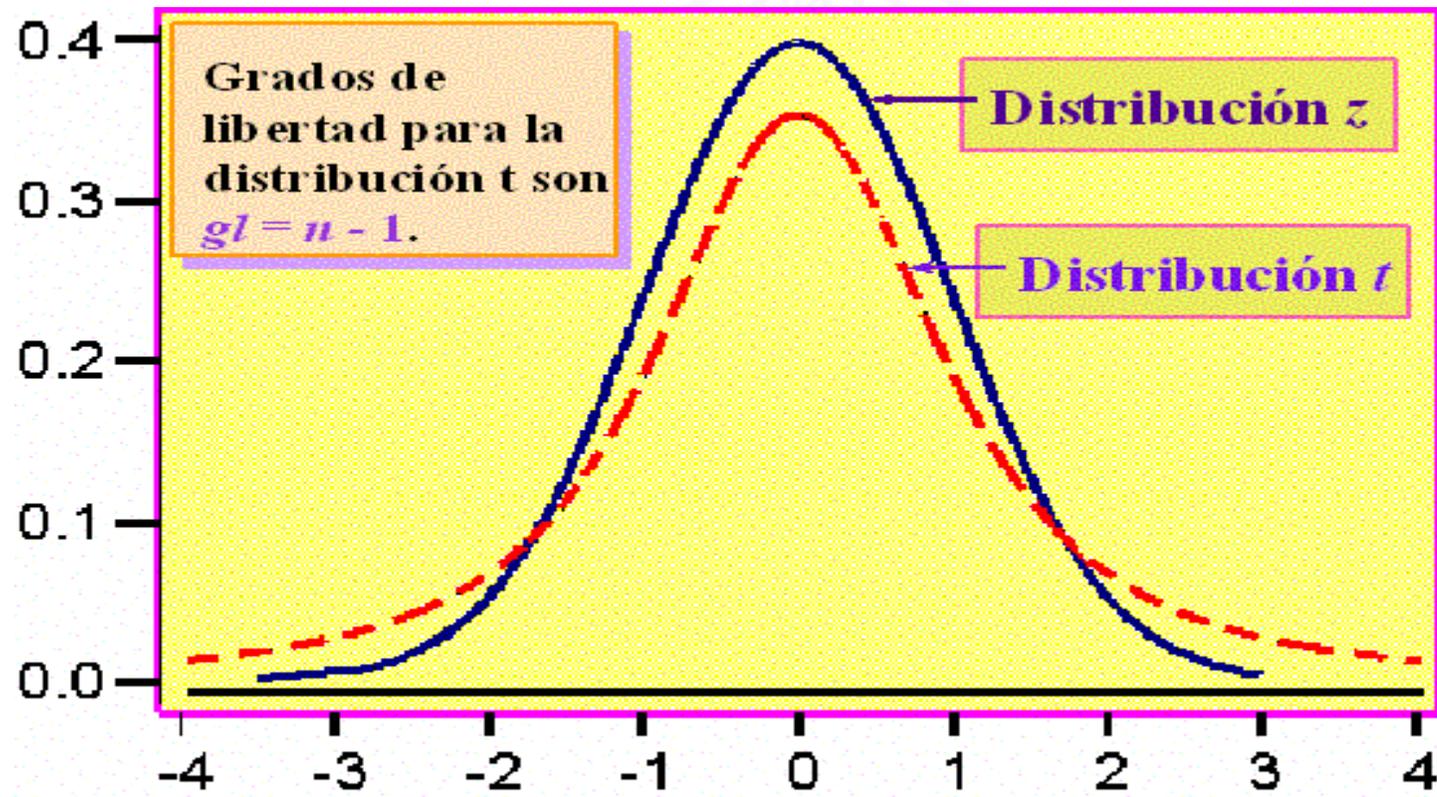
Estadístico T

Al estimar σ a través de S la distribución del Estadístico T es:

$$T = \frac{(\bar{X} - \mu)}{\frac{S}{\sqrt{n}}}$$

El estadístico T, tiene distribución t-Student con $(n-1)$ grados de libertad.

Distribución T-Student y distribución Normal estándar



Algunas características de la distribución t de Student

- Simétrica con respecto a la media cero.
- Función de densidad con forma de Campana y es asintótica al eje de las abscisas
- La desviación estándar de la distribución t-Student dependen del número de grados de libertad.
- Tiene colas más pesadas que la distribución normal
- Para un n "grande" (infinito) la distribución t-Student se aproxima a la distribución normal estándar.

Percentiles de la distribución T-Student

α r	0,25	0,2	0,15	0,1	0,05	0,025
1	1,000	1,376	1,963	3,078	6,314	12,706
2	0,816	1,061	1,386	1,886	2,920	4,303
3	0,765	0,978	1,250	1,638	2,353	3,182
4	0,741	0,941	1,190	1,533	2,132	2,776
5	0,727	0,920	1,156	1,476	2,015	2,571
6	0,718	0,906	1,134	1,440	1,943	2,447
7	0,711	0,896	1,119	1,415	1,895	2,365
8	0,706	0,889	1,108	1,397	1,860	2,306
9	0,703	0,883	1,100	1,383	1,833	2,262
10	0,700	0,879	1,093	1,372	1,812	2,228
11	0,697	0,876	1,088	1,363	1,796	2,201
12	0,695	0,873	1,083	1,356	1,782	2,179
13	0,694	0,870	1,079	1,350	1,771	2,160
14	0,692	0,868	1,076	1,345	1,761	2,145
15	0,691	0,866	1,074	1,341	1,753	2,131
16	0,690	0,865	1,071	1,337	1,746	2,120
17	0,689	0,863	1,069	1,333	1,740	2,110
18	0,688	0,862	1,067	1,330	1,734	2,101
19	0,688	0,861	1,066	1,328	1,729	2,093
20	0,687	0,860	1,064	1,325	1,725	2,086
21	0,686	0,859	1,063	1,323	1,721	2,080
22	0,686	0,858	1,061	1,321	1,717	2,074
23	0,685	0,858	1,060	1,319	1,714	2,069
24	0,685	0,857	1,059	1,318	1,711	2,064
25	0,684	0,856	1,058	1,316	1,708	2,060
26	0,684	0,856	1,058	1,315	1,706	2,056
27	0,684	0,855	1,057	1,314	1,703	2,052
28	0,683	0,855	1,056	1,313	1,701	2,048
29	0,683	0,854	1,055	1,311	1,699	2,045
30	0,683	0,854	1,055	1,310	1,697	2,042
40	0,681	0,851	1,050	1,303	1,684	2,021
60	0,679	0,848	1,045	1,296	1,671	2,000
120	0,677	0,845	1,041	1,289	1,658	1,980

El valor del Percentil 0.975 de La distribución T-Student con 15 g.l es 2.13.

El valor del Percentil 0.975 de la distribución T-Student con 120 g.l es 1.98.

En la medida que los g.l. aumentan, el valor del percentil 0.975 de la distribución T-Student se aproxima a 1.96

Intervalo de confianza $(1 - \alpha)\%$ para una media μ

$$\left(\bar{X} - t_{(n-1), 1-\alpha/2} \frac{S}{\sqrt{n}} ; \bar{X} + t_{(n-1), 1-\alpha/2} \frac{S}{\sqrt{n}} \right)$$

Observación:

La expresión $\frac{S}{\sqrt{n}}$ es el error estándar de la media.

La expresión $t_{(n-1), 1-\alpha/2}$ representa el valor del percentil 0.975 de la distribución T-Student, el cual se aproxima a 1.96 en la medida que el tamaño de muestra crece.

Ejemplo

Se midió colesterol total a 100 pacientes varones adultos de la RM escogidos al azar, los resultados obtenidos arrojan una media de 235 mg/dl y una desviación estándar de 35 mg/dl. Estime el colesterol medio total de la población de varones adultos a través de un intervalo de 95% confianza.

Estimaciones de las Medidas de resumen a considerar:

$$\bar{X} = 235 \quad , \quad S = 35 \quad , \quad n = 100$$

Aplicación directa de la fórmula del intervalo :

$$\text{Limite inferior: } 235 - 2 \times 35 / 10 = 228$$

$$\text{Límite superior: } 235 + 2 \times 35 / 10 = 242$$

Intervalo de 95% de confianza:

$$(228 ; 242)$$

Interpretación:

Con una probabilidad de 0.95 el intervalo (228 ; 242) contiene a la media del colesterol total de varones adultos de la RM.

Observación

Debemos tener en cuenta que el intervalo, $(228, 242)$, tiene una probabilidad de 0.05 de NO contener a la media del Colesterol Total de varones adultos de la RM.

Nota: En general esto no se considera, sin embargo es importante tenerlo en cuenta cuando realizamos investigación científica

Estimación por IC95% para una media utilizando STATA

The screenshot displays the Stata/IC 15.1 software interface. The main menu bar includes 'Archivo', 'Edición', 'Datos', 'Gráficos', 'Estadísticas', 'Usuario', 'Ventana', and 'Ayuda'. The 'Estadísticas' menu is open, showing a list of statistical options. The 'Intervalos de Confianza (ICs)' option is highlighted, and its sub-menu is also open, listing various confidence interval calculators. The 'Calculador de IC normal' option is selected. The command window on the left shows the command `use "C:\Users\Yasna\Downloads\ENS2017_2_ (1).dta"` and the output window on the right shows the message `date site correct.`. On the right side, the 'Variables' panel is visible, listing variables such as `idencuesta`, `fechainiciof1`, `region`, `comuna`, `zona`, and `idsegmento`.

Stata/IC 15.1 - C:\Users\Yasna\Downloads\ENS2017_2_ (1).dta

Archivo Edición Datos Gráficos Estadísticas Usuario Ventana Ayuda

Ventana de revisión

Filtrar comandos aquí

#	Comando	_rc
1	use "C:\Users\Yasna\Downloads\ENS2017_2_ (1).dta"	

Notes:

1.

Checking (contacti
host not
http://ww
unable to

. use "C:
.

Comando

Intervalos de Confianza (ICs)

- Calculador de IC normal
- Calculador de IC de Poisson
- Calculadora de IC para proporciones
- Calculadora de IC para la varianza
- Calculador de IC para la desviación estándar
- Correlaciones y covarianzas
- Correlaciones por pares
- Correlaciones parciales
- Correlación intraclase
- Correlación tetracórica
- Media aritm./geom./armónica
- Graficar medias/medianas por grupos
- Centiles con ICs
- Crear variable de percentiles
- Crear variables que identifican los cuantiles

Sumario y estadísticas descriptivas

- Tablas de frecuencias
- Otras tablas
- Tests de hipótesis clásicos
- Tests de hipótesis no paramétricos
- Gráficos de distribución y tests
- Tests multivariantes de medias, covarianzas y normalidad

date site correct.

Variables

Filtrar variables aquí

Nombre
idencuesta
fechainiciof1
region
comuna
zona
idsegmento

Propiedades

Variables

Nombre
Etiqueta
Typo
Formato
Etiqueta
Notas

Datos

Nombre
ENS20

Estimación por IC95% para una media utilizando STATA

The screenshot shows the STATA interface with the 'ci' dialog box open. The dialog box is titled 'ci - Intervalos de confianza para medias, proporciones y conteos'. It has three tabs: 'Principal', 'by/if/in', and 'Pesos'. The 'Principal' tab is active.

Intervalos de confianza:

- Medias
- Proporciones
- Varianzas
- Desviaciones estándar

Variables: (dejar vacío para todas las variables)
colesterol_total

Opciones:

Tipo de intervalo de confianza:

- Normal
- Poisson (conteos)
- Poisson (tasas)

Variable/exposición: []

Nivel de confianza: 95

Separador de línea cada N variables (poner 0 para ninguno): 5

Agregar salida para todos los grupos combinados (sólo para usar con by)

Buttons: ? R [] OK Cancelar Enviar

Background STATA Interface:

Stata/IC 15.1 - C:\Users\Yasna\Downloads\ENS2017_2_(1).dta

Archivo Edición Datos Gráficos

Ventana de revi... [] [] []

Filtrar comandos aq []

#	Comando	_rc
1	use "C:\Users\Yasna\D...	

Notes:

```
1.
Checking
(contacti
host not
http://ww
unable to
. use "C:
.
```

Comando

Variables:

Nombre
idencuesta
fechainiciof1
region
comuna
zona
idsegmento

Propiedades:

Nombre	ENS20
Etiqueta	
Notas	
Variable	1,162
Observa	6,233
Tamaño	70.64M
Memori	128M

Estimación por IC95% para una media utilizando STATA

Output:

```
. ci means colesterol_total
```

Variable	Obs	Mean	Std. Err.	[95% Conf. Interval]	
colesterol	3,715	181	1	180	183

inta.uchile.cl

