Estimación por Intervalos de confianza

Matemáticas Aplicadas II

¿Qué es estimar por intervalos?

NO buscamos un valor exacto

↓

Buscamos un intervalo donde esté el valor real

Estimación por intervalos = dar un rango de valores probables

Conceptos clave

- **Parámetro** → valor real *(desconocido y que queremos estimar)*
- **Estimador o estadístico** → valor del parámetro en la *muestra*
- **Intervalo de confianza (IC)** → *rango* donde está el parámetro

Ejemplo

Queremos saber el gasto medio mensual en ocio de los adolescentes de 17 años de A Coruña. Para ello tomamos una muestra de 50 jóvenes de esa edad elegidos aleatoriamente.

Tras las encuestas, observamos que, con un 95% de confianza, la media del gasto oscila entre 110 y 130€

❌ No decimos: “la media es 120”

✅ Decimos: “la media de gasto en ocio de los adolescentes está entre 110 y 130” con un nivel de confianza del 95%

Idea clave:

No damos un valor exacto → *damos un intervalo donde está el parámetro.* Damos una aproximación del valor con un error máximo calculado, y al mismo tiempo, calculamos el nivel de confianza para ese margen de error.

Interpretación correcta

(muy importante PAU)

- **CORRECTO:** "Con un **95% de confianza**, (*contextualizar) la media de gasto en ocio de los adolescentes,* está entre 110 y 130"
- **INCORRECTO** "hay una **probabilidad del 95%** de que el parámetro esté en este intervalo"

Teorema central del límite

Es fundamental en Estadística, y dice:

- Si tomas una muestra muy grande (n>30) aunque los datos originales no sean “normales”, y calculas sus medias,
- esas medias tienden a formar una **distribución normal.**

En pocas palabras: **El promedio de muchos datos aleatorios se comporta de manera predecible y “normal”.**

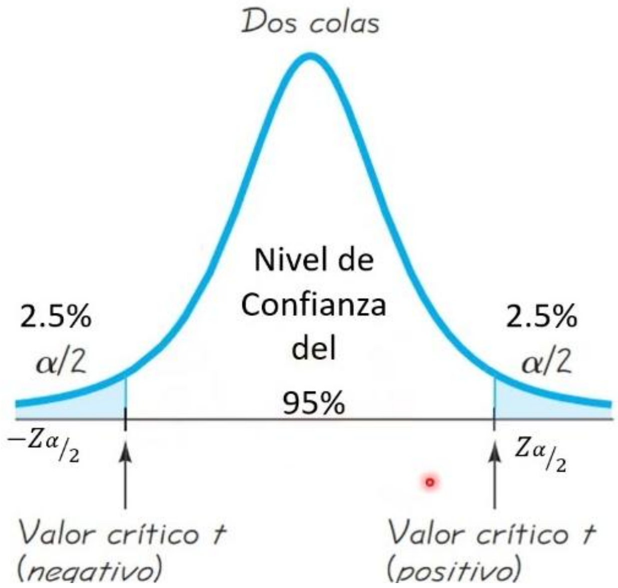
Ejemplo:
Aunque las alturas de unas pocas personas puedan variar mucho, si haces promedios de grupos grandes de personas, verás que se agrupan de forma muy ordenada en torno a un valor central.

Teorema central del límite

- Una consecuencia inmediata del teorema central del límite es que si la muestra es lo bastante grande, la distribución que sigue el estimador o estadístico es normal, con lo que una vez tipificada a una N(0,1) puedo conocer todos sus valores consultando la tabla.

- Así conseguimos trasladarnos desde una distribución de partida cualquiera a una Z: N(0,1) y hacer el estudio con una campana de Gauss

Intervalo de Confianza



Entre sus dos extremos, está el parámetro que se quiere aproximar.

El **nivel de confianza, NC = (1 - α)** es la probabilidad que hay de que esté dentro de este intervalo

El **nivel de significancia, α,** es la probabilidad de quedar fuera de ese intervalo.

Fórmula general

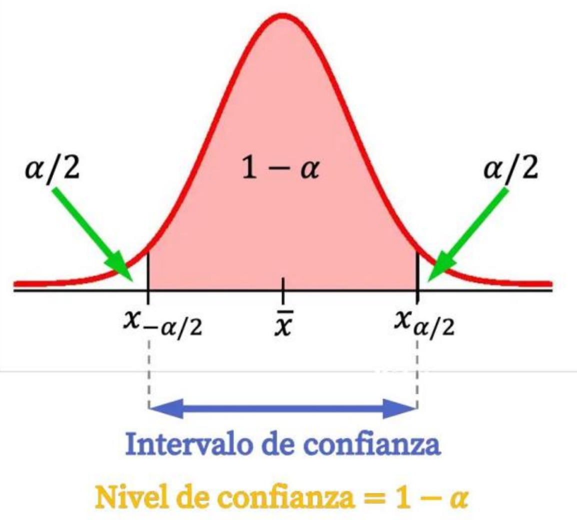
$$IC = estimación \pm error$$

Idea intuitiva

- En el centro del intervalo → lo que estimamos de la muestra
- Radio del intervalo, margen de error → cuánto nos podemos equivocar

Idea Calve

- Nivel de confianza → área bajo la campana
- Buscamos el valor z que deja esa área en el centro



Ejemplo NC del 0,95

¿Qué significa 95%?

Nivel de confianza = **0,95** $\longrightarrow$ Queremos que el **95%** esté en el centro

¿Qué pasa con el resto?

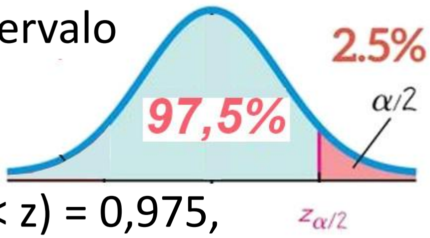
La campana completa vale 1 $\longrightarrow 1 - 0{,}95 = \mathbf{0{,}05} = \boldsymbol{\alpha}$, que es lo que queda **fuera del intervalo,** las 2 colas $\longrightarrow$ hay $\frac{\boldsymbol{\alpha}}{\mathbf{2}} = \mathbf{0{,}025}$ en cada cola

¿Qué buscamos en la tabla?

Las tablas dan: P(Z < z)

Necesitamos el valor acumulado hasta la derecha del intervalo

$$1 - 0{,}025 = 0{,}975 = 97{,}5\%$$



Buscamos en la tabla (*consulta de TABLA INVERSA*) P(Z < z) = 0,975, $\boldsymbol{z = 1{,}96}$

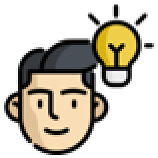
Errores típicos

- No hacer tabla inversa
- Buscar directamente 0,95 en la tabla
- Hay que buscar: (1 + confianza) / 2 = **0,975**

Interpretar mal:

- “probabilidad del 95%”
- “nivel de confianza del 95%”

Resumen



Confianza = 0,95

$\downarrow$

Resto = 0,05

$\downarrow$

Cada cola = 0,025

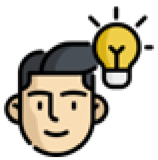
$\downarrow$

Buscamos 1-0,025= 0,975
en tabla inversa

$\downarrow$

z = 1,96

Resumen



El valor crítico $z_{\frac{\alpha}{2}}$ se obtiene dejando en el centro el nivel de confianza y repartiendo el resto en las dos colas de la distribución normal.

Tipos de ejercicios

Dependiendo del parámetro que se desea estimar, son dos:

I. **Media**
II. **Proporción**

Intervalo de confianza para la media poblacional μ con desviación típica conocida

Matemáticas Aplicadas II

Fórmulas

Siempre con desviación típica conocida

$$I.C. = (\bar{X} - E, \bar{X} + E)$$

$$E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

μ está dentro de este intervalo con un nivel de confianza $1 - \alpha$, o lo que es lo mismo:

$$P(\bar{X} - E < \mu < \bar{X} + E) = 1 - \alpha$$

Tamaño mínimo de la muestra ***n (entero)*** lo despejamos directamente de la fórmula del error máximo

Qué significa cada cosa

$\bar{X} \rightarrow$ media muestral

$z \rightarrow$ valor de la tabla, una vez tipificada

$\sigma \rightarrow$ desviación típica

$n \rightarrow$ tamaño muestra

$E \rightarrow$ Error máximo o radio del I.C.

EXAMPLE

En un hospital se ha tomado la temperatura a una muestra de 64 pacientes, para estimar la temperatura media de sus enfermos.

La media de la muestra ha sido de 37,1°C y la desviación típica de la población de 1,04°C.

Calcula un intervalo de confianza para la media poblacional con un nivel de confianza del 99%

Paso 1: nos fijamos en los datos, piden un IC para la *media*

- **media** = 37,1°C
- σ = 1,04°C
- n = 64
- confianza = 99%

Paso 2: valor de z

$$1-\alpha=0{,}99\rightarrow\alpha=0{,}01\rightarrow\frac{\alpha}{2}=0{,}005\rightarrow1-0{,}005=0{,}995$$

Hacemos uso de tabla inversa para buscar $P(Z<z_{0{,}005})=0{,}995$

$\rightarrow Z_{0{,}005}=2{,}58$

EXAMPLE

Paso 3: error $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$E = 2{,}58 \cdot \frac{1{,}04}{\sqrt{64}} = 0{,}3354$$

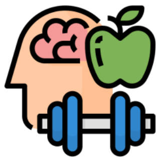
Paso 4: intervalo

$$IC = 37{,}1 \pm 0{,}3354 \rightarrow (36{,}76;\ 37{,}44)$$

Interpretación

Con un 99% de confianza, la temperatura media, de los pacientes del hospital, está entre 36,76°C y 37,74°C

Ejercicio competencial



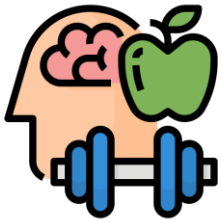
Una de las principales novedades de las pruebas PAU 2025 fue que el examen de cada materia incluyó un ejercicio obligatorio y de carácter "más competencial". Aunque las notas se hacen públicas la semana siguiente de realizarse el examen, los miembros del grupo de trabajo de la materia Matemáticas Aplicadas a las Ciencias Sociales II estaban interesados en determinar cuanto antes si se habían producido cambios relevantes en la nota media de la materia que coordinan con relación a las notas de cursos pasados. Con este objetivo contactaron previamente con un grupo de correctores, de los que cada uno de ellos se comprometió a corregir un máximo de 25 exámenes el primer día. Por los datos de otros cursos, las notas de esta materia pueden suponerse que siguen una distribución normal con desviación típica igual a 1,5.

1.1. Si se quiere estimar esta nota media con un error máximo de 0,25, empleando un nivel de confianza del 95%, ¿cuál es el número mínimo de correctores que se necesitan?
1.2. Una vez corregidos los 100 primeros exámenes, la nota media resultó ser igual a 7,2. A partir de esta muestra, calcule un intervalo de confianza con nivel de confianza del 95% de la nota media.

Contextualice la respuesta obtenida.

Nota: Para resolver algunos de los apartados anteriores pueden emplearse algunos de los siguientes valores relacionados con las tablas de la normal estándar: $P(|Z|<1)=0{,}6826$; $P(Z<2)=0{,}9772$; $P(Z>0{,}5)=0{,}3085$; $P(Z>1{,}96)=0{,}025$.

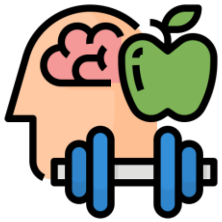
Ejercicio 2



El salario (en €) de los trabajadores de una empresa se distribuye normalmente con desviación típica σ=300€. Se preguntó a 36 trabajadores elegidos al azar, y se establece que el salario medio de los trabajadores de la empresa oscila entre 1552€ e 1748€

¿Cuál ha sido el salario medio de los trabajadores de la muestra? ¿Con qué nivel de confianza se ha establecido el intervalo anterior?

Ejercicio 3



2022 Ordinaria

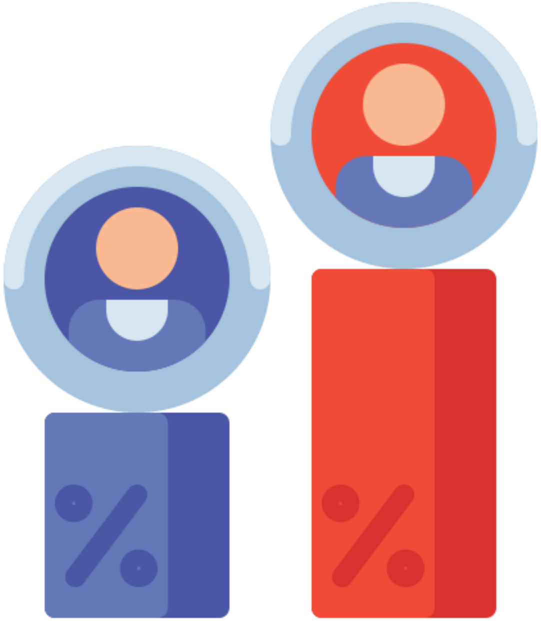
Tomamos una muestra aleatoria de 36 facturas de consumo mensual de luz (en euros) y el intervalo de confianza obtenido al 95% para el consumo mensual medio es [60.1, 69.9]. Según esta información:

a) ¿Cuál fue el consumo medio muestral de luz?

b) ¿Cuál es el error máximo cometido?

c) Determine un intervalo de confianza al 90% para el consumo medio de luz *(para resolver este apartado ANTES hay que despejar en el b) el valor de la desviación típica)*

Intervalo de confianza para la proporción p



Matemáticas Aplicadas II

Fórmulas

$$I.C. = (\hat{p} - E, \hat{p} + E)$$

$$E = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

p está dentro de este intervalo con un nivel de confianza 1 – α, o lo que es lo mismo:

$$P(\hat{p} - E < p < \hat{p} + E) = 1 - \alpha$$

Tamaño mínimo de la muestra ***n (entero)*** $n \geq \left(\frac{Z_{\alpha/2}}{E}\right)^2 \cdot (\hat{p} \cdot \hat{q})$

Utilidades

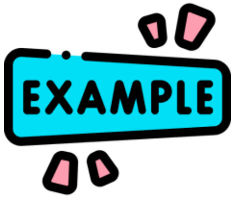
- Encuestas
- Estimación de porcentajes
- Estudios sociales

Proceso paso a paso

1) Identificar tipo (media o proporción)
2) Elegir fórmula
3) Calcular error
4) Construir intervalo
5) Interpretar

Idea Clave para interpretar

- Más confianza → intervalo más grande
- Menos error → muestra más grande
- Más muestra → más precisión



El ministerio de comunicación quiere estimar un porcentaje de personas que utilizan las redes sociales, durante su jornada laboral. Para ello realizó una encuesta a 400 personas a la salida del trabajo y 240 de ellas afirmaron haberlas consultado.

Halla el intervalo de confianza para el porcentaje de usuarios con un nivel de confianza del 95%

Datos: n = 400, 240 usan redes, NC = 0,95

Paso 1: proporción $\hat{p} = 240/400 = 0{,}6$

Paso 2: fórmula $E = Z_{\alpha/2} \cdot \sqrt{\dfrac{\widehat{p} \cdot \widehat{q}}{n}}$

Paso 3: error

$$E = 1{,}96\sqrt{\frac{0{,}6 \cdot 0{,}4}{400}} = 0{,}048$$

Paso 4: intervalo

$$(0{,}552;\ 0{,}648)$$

Interpretación

Con un **95% de confianza**, la proporción de personas que utilizan las redes sociales durante su jornada laboral, **está entre 0,552 y 0,648.**

Errores típicos

- No dividir entre $\sqrt{400}$
- Confundir media y proporción
- No interpretar
- Interpretar mal el 95%

EXAMPLE

En las últimas elecciones se ha tomado una muestra de 450 personas a la salida de los colegios electorales y 125 de ellas afirmaron haber votado al partido *A*.

Halla el intervalo de confianza para el porcentaje de votantes de *A* con un nivel de confianza del 90%

EXAMPLE

$$125\ personas\ de\ 450\ \rightarrow \boxed{\hat{p} = \frac{125}{450} = 0{,}28}$$

$$1 - \alpha = 0{,}90 \rightarrow \alpha = 0{,}10 \rightarrow \frac{\alpha}{2} = 0.05 \rightarrow 1 - 0.05 = 0.95$$

Ahora hacemos uso de **tabla inversa**, busco un valor $\mathrm{P}(\mathrm{Z} < \mathrm{Z}_{0.05}) = 0.95$

El valor exacto no está en la tabla, así que hago un promedio entre los valores más próximos: 1.64 y 1.65

que es 1.645 por tanto

$$\boxed{\mathrm{Z}_{0.05} = 1.645}$$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515

EXAMPLE

Calculamos el error para el intervalo, la fórmula correcta es:

$$E = Z_{\alpha/2} \cdot \sqrt{\frac{\widehat{p} \cdot \widehat{q}}{n}}$$

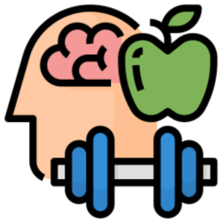
$$E = 1.645 \cdot \sqrt{\frac{0{,}28 \cdot 0{,}72}{450}} \approx 0{,}021$$

$$estimación \pm error \rightarrow 0{,}28 \pm 0{,}021 \rightarrow \pm \begin{cases} 0{,}28 - 0{,}021 \approx 0{,}26 \\ 0{,}28 + 0{,}021 \approx 0{,}301 \end{cases}$$

IC=(0,26; 0,301)

Con un 90% de confianza, la proporción de votantes del partido A, está entre el 0,26 y 0,301

Ejercicio 1



De una muestra de 500 personas, 325 tienen teléfono móvil. Determina un intervalo de confianza para estimar la proporción de usuarios de teléfono móvil de la población con un nivel de confianza:

a) Del 95%

b) Del 99%.

$$a) \quad 325 \; de \; 500 \rightarrow \hat{p} = \frac{325}{500} = 0{,}65 \rightarrow \hat{q} = 1 - 0{,}65 = 0{,}35$$

$$1 - \alpha = 0{,}95 \rightarrow \alpha = 0{,}05 \rightarrow \frac{\alpha}{2} = 0{,}025 \rightarrow \mathrm{Z}_{0{,}025} = 1{,}96$$

$$E = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} \qquad E = 1{,}96 \cdot \sqrt{\frac{0{,}65 \cdot 0{,}35}{500}} \approx 0{,}0418$$

$$extremo\ inferior \rightarrow 0{,}65 - 0{,}0418 = 0{,}6082 \approx 0{,}61$$

$$extremo\ superior \rightarrow 0{,}65 + 0{,}0418 = 0{,}6918 \approx 0{,}69$$

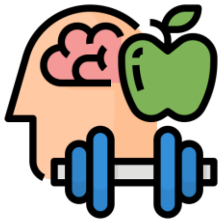
$$IC = (0{,}61, 0{,}69)$$

Con un 95% de confianza, la proporción de personas que tiene teléfono móvil, está entre el 61% y el 69%

Resuelve tú el apartado b)
solución $IC = (0{,}59;\ 0{,}71)$

Interpretación, para estar más seguros → abrimos el intervalo

Ejercicio 2



De una muestra de 300 bombillas, 24 han resultado ser defectuosas.

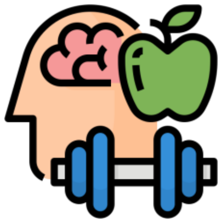
Al determinar el intervalo de confianza para la proporción de bombillas defectuosas de la población con un nivel de significación de 0,01 , ¿cuál es el error máximo admisible?

24 bombillas defectuosas de 300 $\rightarrow \hat{p} = \dfrac{24}{300} = 0,08$

$\alpha = 0{,}01 \rightarrow z_{\alpha/2} = z_{0{,}005} = 2{,}58$

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}} = 2{,}58\cdot\sqrt{\frac{0{,}08\cdot 0{,}92}{300}} = 0{,}04$$

Ejercicio 3



De una muestra de 1.000 habitantes, 650 leen cierto periódico. ¿Cuál debe ser el tamaño mínimo de una muestra, si se pretende que el error máximo admisible cometido, al estimar la proporción poblacional, sea inferior al 5% , para un nivel de confianza del 90%?

$$n \geq \left(\frac{Z_{\alpha/2}}{E} \right)^2 \cdot (\hat{p} \cdot \hat{q})$$

650 habitantes leen cierto periódico de 1.000 $\rightarrow \hat{p} = \dfrac{650}{1.000} = 0{,}65$

$1 - \alpha = 0{,}90 \rightarrow \alpha = 0{,}1 \rightarrow z_{\alpha/2} = z_{0{,}05} = 1{,}645$

$$n = \frac{1{,}645^2 \cdot 0{,}65 \cdot 0{,}35}{0{,}05^2} = 246{,}25$$

Para que se cumplan las condiciones, y como n tiene que ser un número entero, debe ser mayor o igual que 247.

Resumen

Estimación puntual → un único valor aproximado

Intervalo de Confianza → rango con un error máximo, en el que puede estar el parámetro, con un nivel de confianza preestablecido.

- **IC =** estimación ± error
- Más muestra → más precisión
- Más confianza → intervalo mayor

Esquema mental

Problema

↓

Identificar variable

↓

¿Media o proporción?

↓

¿Intervalo o probabilidad?

↓

Aplicar fórmula

↓

Interpretar en el contexto

Resumen Fórmulas

	Parámetro a estimar		Distribución original de X	Estimador o estadístico		Distribución del estimador	Tipificación N(0,1)	*n* tamaño de la muestra	*N.C.* Nivel de confianza	Valor crítico	*E* Error	Intervalo de confianza	Tamaño de la muestra
Estimación puntual	$\boldsymbol{\mu}$	media poblacional	$N(\mu,\sigma)$	$\bar{X}$	media muestral	$N\left(\mu,\frac{\sigma}{\sqrt{n}}\right)$	$Z=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$	cualquiera	No procede				
			Desconocida, pero con σ conocida					$n \geq 30$					
	$\boldsymbol{p}$	proporción poblacional	Binomial	$\hat{p}$	proporción muestral	$N\left(p,\sqrt{\frac{p\cdot q}{n}}\right)$	$Z=\frac{\hat{p}-p}{\sqrt{\frac{p\cdot q}{n}}}$	$n \geq 30$					
Estimación por intervalos	$\boldsymbol{\mu}$	media poblacional	$N(\mu,\sigma)$	$\bar{X}$	media muestral	$N\left(\mu,\frac{\sigma}{\sqrt{n}}\right)$	$Z=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$	cualquiera	$1-\alpha$	$z_{\frac{\alpha}{2}}$ tabla inversa → $P\left(Z<z_{\frac{\alpha}{2}}\right)=\frac{1+NC}{2}$	$z_{\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$	$\bar{X}\pm E$	$n=\frac{z_{\frac{\alpha}{2}}^2\cdot\sigma^2}{E^2}$
			Desconocida, pero con σ conocida					$n \geq 30$					
	$\boldsymbol{p}$	proporción poblacional	Binomial	$\hat{p}$	proporción muestral	$N\left(p,\sqrt{\frac{p\cdot q}{n}}\right)$	$Z=\frac{\hat{p}-p}{\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}}$	$n \geq 30$	$1-\alpha$	$z_{\frac{\alpha}{2}}$ tabla inversa → $P\left(Z<z_{\frac{\alpha}{2}}\right)=\frac{1+NC}{2}$	$z_{\frac{\alpha}{2}}\cdot\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$	$\hat{p}\pm E$	$n=\frac{z_{\frac{\alpha}{2}}^2\hat{p}\hat{q}}{E^2}$

Matemáticas Aplicadas II

Enfoque CIUGA (lo que realmente piden)

En PAU Galicia se evalúa:

- ✔ saber **identificar el modelo**
- ✔ aplicar **fórmulas correctamente**
- ✔ hacer **cálculos con normal**
- ✔ interpretar resultados
- ✔ entender **error y confianza**

Errores típicos de los alumnos

- Confunden:
  - parámetro vs estadístico
  - media vs proporción
  - probabilidad vs intervalo
- Olvidan la interpretación final, o hacen interpretación errónea