Apuntes de Probabilidad y Estadística II

Leonardo H. Añez Vladimirovna[1]

Universidad Autónoma Gabriél René Moreno,
Facultad de Ingeniería en Ciencias de la Computación y Telecomunicaciones,
Santa Cruz de la Sierra, Bolivia

20 de marzo de 2020

[1]Correo Electrónico: toborochi98@outlook.com

Notas del Autor

Estos apuntes fueron realizados durante mis clases en la materia `MAT305 (Probabilidad y Estadística II)`, acompañados de referencias de libros, fuentes y código que use a lo largo del curso, en el período `I-2018` en la Facultad de Ingeniería en Ciencias de la Computación y Telecomunicaciones.

Para cualquier cambio, observación y/o sugerencia pueden enviarme un mensaje al siguiente correo:

`toborochi98@outlook.com`

Índice general

Capítulo 1

Variables Aleatorias

Una variable aleatoria x (desde ahora denotada por **v.a.**) es una función definida sobre el espacio muestral S con valores en $\mathbb{R}$ que a cada elemento de S (Punto muestral) hace corresponder un número real $x = X$.

$$x = X(w) \in Rec_X \subseteq \mathbb{R}$$

Gráficamente

Notación Conjuntista

$$X = \{(w, x)\ w \in S, x = X(w) \in \mathbb{R}\} \subseteq S \times \mathbb{R}$$

Donde:

- S: Conjunto Partida (Espacio Muestral).
- $\mathbb{R}$: Conjunto de llegada.
- w: Elemento de S (Punto Muestral).
- x: Valor de la **v.a.** X.
- Rec_X: Recorrido de X.
- X: Función **v.a.** (Conjunto de Pares Ordenados).

Notaciones

Las **v.a.** se denotan con letras mayúsculas tales como X, Y o Z, y los valores correspondientes con letras minúsculas.

1.1. Clasificación de Variables Aleatorias

- **Discreta:** Cuyo recorrido es un conjunto finito o infinito numerable de valores:

$$X \text{ es } \textbf{v.a.} \text{ discreta} \Rightarrow \begin{cases} \text{Conjunto Finito de Valores} \\ \text{Conjunto Infinito Numerable de Valores} \end{cases}$$

- **Contínua:** Es aquella cuyo recorrido es conjunto finito no numerable de valores, puede tomar cualquier valor en un intervalo o conjunto.

♦ En general las **v.a.** *discretas* representan datos que provienen del *conteo* de número de elementos. Pueden ser número de titulados, número de estudiantes, etc. Mientras que las **v.a.** *contínuas* representan mediciones, como longitud, capacidad,etc.

1.2. Función de Probabilidad de una Variable Aleatoria

También llamada función de cuantía o función de masa de probabilidad de una **v.a.**.

Se denomina función de probabilidad de una **v.a.** discreta X a una función p o f, cuyo valor es $p(x)$ o $P(X = x)0$ ya que a cada valor distinto de la **v.a.** discreta X hace corresponder en un número entre los valores $[0, 1]$ que es su probabilidad, de ahí el nombre de función de cuantía o función de probabilidad. Estos valores satisfacen las siguientes condiciones:

1. $P(x) \geq 0 \quad ; \quad \forall x \in \mathbb{R}$
2. $\displaystyle\sum_{x_i \in Rec_X} p(x_i) = 1$
    - Si $Rec_X = \{x_1, x_2, \ldots, x_n\}$ entonces la condición **(II)** es: $\displaystyle\sum_{i=1}^{n} p(x_i) = 1$
    - Si $Rec_X = \{x_1, x_2, \ldots, x_n, \ldots\}$ entonces la condición **(II)** es: $\displaystyle\sum_{i=1}^{\infty} p(x_i) = 1$

Si A es un evento en el recorrido de la **v.a.** discreta X entonces la probabilidad de A es el número:

$$P(A) = \sum P(X = x) = \sum p(x)$$

Nota:

$$P(X = x) \begin{cases} p(x) \geq 0; & \forall x \in \mathbb{R} \\ f(x) \geq 0; & \forall x \in \mathbb{R} \end{cases}$$

La función de probabilidad de una **v.a.** discreta X se puede expresar por:

- **Un Conjunto:**

$$p = \{(x, P(X))/x \in D_p\}$$

- **Una Tabla:**

x_i	x_1	x_2	$\ldots$	x_n
$p(x_i)$	$p(x_1)$	$p(x_2)$	$\ldots$	$p(x_n)$

- **Una Gráfica:**

1.3. Función de Distribución Acumulada (FDA)

El valor de la **FDA** de una **v.a.** discreta X, que es $F(x)$, viene dada por la sumatoria de las probabilidades, desde un valor mínimo t hasta un valor específico x; esto es:

$$F(x) = P(X \leq x) = \sum_{t \leq x} P(t), \quad \forall x \in \mathbb{R}$$

1.3.1. Representación Gráfica

Valores $F(x)$ aumentan en *saltos*, presentando entonces la forma de una escalera:

1.3.2. Caso Continuo

Función de Densidad

f es función densidad, si $f(x)$ cumple las siguientes condiciones:

(I) $f(x) \geq 0; \quad \forall x \in \mathbb{R}$

(II) $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$

(III) $\displaystyle p(a \leq x \leq b) = \int_{a}^{b} f(x)dx$

FDA

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$$

1.3.3. Propiedades de la FDA

Caso Discreto

1. $0 \leq F(x) \leq 1; \quad \forall x \in \mathbb{R}$
2. $F(-\infty) = 0$
3. $F(+\infty) = 1$
4. $P(X \leq a) = F(a)$
5. $P(X > a) = 1 - P(X < a) = 1 - F(a)$
6. $P(X < a) = \begin{cases} F(a-1), \text{si } a \in \mathbb{Z} \\ F([\![a]\!]), \text{si } a \notin \mathbb{Z} \end{cases}$
7. $P(X \leq -a) = 1 - P(X \leq a) = 1 - F(a)$
8. $P(a < X \leq b) = F(b) - F(a)$
9. $P(a \leq X \leq b) = F(b) - F(a) + P(X = x)$
10. $P(a < X < b) = F(b) - F(a) - P(X = b)$
11. $P(X = x_i) = F(x_i) - F(x_{i-1})$

Caso Continuo

1. $0 \leq F(x) \leq 1; \quad \forall x \in \mathbb{R}$
2. $F(-\infty) = 0$
3. $F(+\infty) = 1$
4. $P(X \leq a) = P(X < a) = F(a)$
5. $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$
6. $P(X \geq a) = 1 - P(X < a) = 1 - F(a)$
7. $P(X \leq -a) = 1 - P(X \leq a) = 1 - F(a)$
8. $P(a < X \leq b) = P(a \leq X \leq b) = P(a < X < b)$
9. $f(x) = \dfrac{dF(x)}{dx}$

1.4. Esperanza Matemática

Sea X una **v.a.** con función de probabilidad f definida por $f(x)$. La *esperanza matemática* de X, denotada por $E(x), \mu$ ó μ_x; está dada por:

$$E(x) = \mu = \mu_x = \begin{cases} \displaystyle\sum_x x \cdot p(x), \ \text{Si } X \text{ es } \textbf{v.a.} \text{ Discreta} \\ \displaystyle\int_{-\infty}^{+\infty} x \cdot f(x)dx, \ \text{Si } X \text{ es } \textbf{v.a.} \text{ Continua} \end{cases}$$

Propiedades

1. $E(a) = a$
2. $E(x \pm a) = E(x) \pm a$
3. $E(ax) = aE(x)$
4. $E(ax \pm b) = aE(x) \pm b$

1.4.1. Varianza

Notaciones: $V(x), \sigma^2, \sigma_x^2$

$$V(x) = \sigma^2 = \begin{cases} E[x-\mu]^2 = \sum_x (x-\mu)^2 f(x); \text{ Si } X \text{ es } \textbf{v.a.} \text{ Discreta} \\ E[x-\mu]^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx; \text{Si } X \text{ es } \textbf{v.a.} \text{ Continua} \end{cases}$$

Propiedades

1. $V(x) \geq 0$
2. $V(a) = 0$
3. $V(ax) = aV(x)$
4. $V(ax \pm b) = a^2 V(x)$
5. $V(x) = E(x^2) - [E(x)]^2$

1.5. Función de Probabilidad Conjunta

1.5.1. Función de Cuantía Conjunta

$$P(X = x, Y = y) = f(x, y) = p(x, y)$$

es el valor de una función de cuantía conjunta de la **v.a.'s** X y Y si:

(I) $f(x, y) = p(x, y) \geq 0$ Para cualquier (x, y) de su dominio.

(II) $\sum_x \sum_y f(x, y) = 1$

(III) $P((x, y) \in A) = \sum_A \sum f(x, y)$ Para cualquier región A del plano XY.

1.5.2. Función de Densidad Conjunta

$$P(X = x, Y = y) = f(x, y) = p(x, y)$$

es el valor de una función de cuantía conjunta de la **v.a.'s** X y Y si:

(I) $f(x, y) = p(x, y) \geq 0$ Para cualquier (x, y) de su dominio.

(II) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1$

(III) $P((x, y) \in A) = \iint_A f(x, y)\, dx\, dy$ Para cualquier región A del plano XY.

1.6. Distribuciones Marginales

Sean X y Y **v.a.** con función de probabilidad conjunta definida por $f(x,y)$. La distribución marginal está dada por:

$$\text{Caso Discreto: }\begin{cases}\text{Distribución Marginal X: } g(x) = \sum_y f(x,y) \\ \\ \text{Distribución Marginal Y: } h(y) = \sum_x f(x,y)\end{cases}$$

$$\text{Caso Continuo: }\begin{cases}\text{Distribución Marginal X: } g(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy \\ \\ \text{Distribución Marginal Y: } h(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx\end{cases}$$

1.6.1. Independencia Estadística de las v.a. X y Y

Sean X y Y **v.a.** discretas o contínuas con función de probabilidad conjunta definida por $f(x,y)$ y distribuciones marginales $g(x)$ y $h(y)$, respectivamente. Se dice que las **v.a.** X y Y serán estadísticamente independientes ssi:

$$f(x,y) = g(x)h(y)$$

Para cualquier (x,y) dentro de sus recorridos.

1.6.2. Esperanza Matemática

Sean X y Y **v.a.** con función de probabilidad definida por $f(x,y)$. La media o esperanza matemática de $g(x,y)$ está dada por:

$$\text{Caso Discreto: } E(g(x,y)) = \mu_{g(x,y)} = \sum_x \sum_y g(x,y) \cdot f(x,y)$$

$$\text{Caso Contínuo: } E(g(x,y)) = \mu_{g(x,y)} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) \cdot f(x,y)\,dx\,dy$$

1.6.3. Covarianza

Mide el grado de relación o asociación de dos variables.

1.6.4. Resultado Importantes

1. $E(x \pm y) = E(x) \pm E(y)$
2. $Cov(x,y) = E(x \cdot y) - E(x)E(y)$
3. $V(x \pm y) = V(x) + V(y) \pm 2Cov(x,y)$
4. $V(ax \pm by) = a^2V(x) + b^2V(y) \pm 2abCov(x,y)$
5. Si X y Y son estadísticamente independientes entonces:
   a) $E(x,y) = E(x)E(y)$
   b) $Cov(x,y) = 0$
   c) $V(x \pm y) = V(x) + V(y)$
   d) $V(ax \pm by) = a^2V(x) + b^2V(y)$
6. Si $x1, x2, x3, \ldots$ son variables aleatorias independientes 2 a 2:

   $$V\left(\sum_{i=1}^{m} x_i\right) = \sum_{i=1}^{m} V(x_i)$$

Capítulo 2

Modelos de Distribución de Probabilidad

2.1. Distribuciones Discretas

2.1.1. Distribución de Bernoulli

Si la probabilidad de que ocurra un evento p y la probabilidad de que no ocurra es $q(q = 1 - p)$, entonces se dice que la **v.a.** discreta X se distribuye según Bernoulli, cuya función de cuantía está dada por:

$$p(x) = f(x) = P(X = x) = \begin{cases} p^x \cdot (1-p)^{1-x} & ; x = 0; 1 \\ 0 & ; \text{otro caso} \end{cases}$$

Y cuya función de distribución acumulada es:

$$F(X) = P(X = x) = \begin{cases} 0 & ; x < 0 \\ q = 1 - p & ; 0 \leq x < 1 \\ 1 & ; x \geq 1 \end{cases}$$

$$X \sim Ber(x; p) \Rightarrow \begin{cases} E(x) = \mu = p \\ V(x) = \sigma^2 = p \cdot q \\ D(x) = \sigma = \sqrt{p \cdot q} \end{cases}$$

Conocida como prueba o ensayo de Bernoulli, es un experimento que solo tiene 2 resultados posibles, a los cuales se los llama:

- Éxito (p)
- Fracaso (q)

2.1.2. Distribución Binomial

Una **v.a.** discreta X tiene distribución lineal si su función de cuantía está dada por:

$$f(x) = \begin{cases} \binom{n}{x} p^x \cdot q^{n-x} & ; x = 0, 1, 2, \ldots, n \\ 0 & ; \text{otro caso} \end{cases}$$

Donde:

- p : Probabilidad de éxito.
- n : Número de ensayo.
- x : Número de éxitos.

Las Funciones de Distribución Acumulada Binomialmente está definida por:

$$F(x) = P(X \leq x) = \begin{cases} 0 & ; x < 0 \\ \sum_{k=0}^{[\![x]\!]} \binom{n}{k} p^k q^{n-k} & ; 0 \leq x < n \\ 1 & ; x \geq n \end{cases}$$

$$X \sim b(x; n, p) \Rightarrow \begin{cases} E(x) = \mu = n \cdot p \\ V(x) = \sigma^2 = n \cdot p \cdot q \\ D(x) = \sigma = \sqrt{n \cdot p \cdot q} \end{cases}$$

Características de la Distribución Binomial

1. Se realiza n pruebas , cada una independiente.
2. p es la probabilidad de éxito en cada prueba que ocurra en un evento y se mantiene constante a travez de las n pruebas.
3. El experimento es con reposición (sustitución por reemplazo).
4. Se da el valor de la **v.a.** X. La variación de x es desde 0 hasta n.

Observaciones

$$f(x) = b(x; n, p) \qquad n \text{ y } p \text{ son parámetros.}$$

Manejo de la Tabla Binomial

1. $\begin{pmatrix} p \leq 0{,}50 \\ n \leq 20 \end{pmatrix} \Rightarrow b(x; n, p) = B(x; n, p) - B(x-1; n, p)$
2. $\begin{pmatrix} p > 0{,}50 \\ n \geq 20 \end{pmatrix} \Rightarrow \begin{cases} \textbf{(i.)}\ b(x; n, p) = B(n-x; n, 1-p) - B(n-x-1; n, 1-p) \\ \textbf{(ii.)}\ b(x; n, p) = b(n-x; n, 1-p) \text{ luego de usar } \textbf{(i.)} \\ \textbf{(iii.)}\ B(x; n, p) = 1 - B(n-x-1; n, 1-p) \end{cases}$

2.1.3. Distribución de Poisson

Una **v.a.** discreta X tiene distribución de Poisson, si su función de cuantía está dada por:

$$p(x) = f(x) = \begin{cases} \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}; & x = 0, 1, 2, \ldots \\ 0; & \text{otro caso} \end{cases}$$

Parámetro: $\lambda > 0$

La distribución de Poisson se obtiene de 2 maneras:

1. $\lim_{\substack{n \to \infty \\ p \to 0}} \binom{n}{x} \cdot p^x \cdot q^{n-x} \backsimeq \dfrac{e^{-\lambda} \cdot \lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$
2. $p(x) = f(x) = \dfrac{e^{-\lambda \cdot t} (\lambda \cdot t)^x}{x!} \qquad x = 0, 1, 2, \ldots$

donde t es la cantidad de medida (intervalo de tiempo, longitud, área, etc...) La **F.D.A.** de Poisson está dada por:

$$F(x) = \begin{cases} 0 & ; x < 0 \\ \sum_{k=0}^{[\![x]\!]} \dfrac{e^{-\lambda} \cdot \lambda^x}{x!} & ; x \geq 0 \end{cases}$$

$$X \sim Poisson(x; \lambda) \Rightarrow \begin{cases} E(x) = \mu = \lambda = n \cdot p \\ V(x) = \sigma^2 = n \cdot p \\ D(x) = \sigma = \sqrt{\lambda} \end{cases}$$

2.1.4. Distribución Geométrica

También llamada Distribución de Pascal, es un modelo útil para aquellos procesos en los que se repiten pruebas hasta llegar al éxito o a un resultado deseado y tiene interesantes aplicaciones en los muestreos realizados de esta manera. Su función de cuantía está dada por:

$$p(x) = f(x) = (1-p)^x \cdot p \qquad x = 0, 1, 2, \ldots \quad ; \quad (0 < p < 1)$$

También implica la existencia de una dicotomía de posibles resultados y la independencia de las pruebas entre sí.

Características de la Distribución Geométrica

1. El proceso consta de un número no definido de pruebas o experimentos separados o separables. El proceso concluirá cuando se obtenga por primera vez el resultado deseado (éxito).
2. Cada prueba puede dar dos resultados mutuamente excluyentes : A y no A
3. La probabilidad de obtener un resultado A en cada prueba es p y la de obtener un resultado no A es q siendo $(p + q = 1)$.

La **F.D.A.** de esta distribución es:

$$F(x) = \begin{cases} 0 & ; x < 0 \\ \displaystyle\sum_{k=0}^{[\![x]\!]} (1-p)^k p & ; x \geq 0 \end{cases}$$

Y se notará de la siguiente manera:

$$X \sim G(p) \Rightarrow \begin{cases} E(x) = \mu = \dfrac{1}{p} \\[2ex] V(x) = \sigma^2 = \dfrac{1-p}{p^2} \\[2ex] D(x) = \sigma = \dfrac{\sqrt{1-p}}{p} \end{cases}$$

2.1.5. Distribución Binomial Negativa

Consideremos ahora un experimento aleatorio consistente en repeticiones independientes de ensayos de Bernoulli con probabilidad de éxito constante, hasta que aparezca el éxito k-ésimo. Es decir, en lugar de fijar el número de ensayos y observar el número de éxitos en esas n realizaciones, se repiten las realizaciones hasta obtener un número determinado de éxitos y contabilizamos los fracasos. Definimos la variable aleatoria con distribución binomial negativa como aquella que modeliza el número de fracasos antes de que aparezca el éxito k-ésimo.

$$p(x) = f(x) = \binom{x + \lambda - 1}{x} (1-p)^x p^\lambda$$

Su **F.D.A.** está dada por:

$$F(x) = \begin{cases} 0 & ; x < 0 \\ \displaystyle\sum_{k=0}^{[\![x]\!]} \binom{k + \lambda - 1}{x} (1-p)^k p^\lambda & ; x \geq 0 \end{cases}$$

Denotada por:

$$X \sim BN(\lambda, p) \Rightarrow \begin{cases} E(x) = \mu = \dfrac{\lambda(1-p)}{p} \\[2ex] V(x) = \sigma^2 = \dfrac{\lambda(1-p)}{p^2} \\[2ex] D(x) = \sigma = \dfrac{\sqrt{\lambda(1-p)}}{p} \end{cases}$$

2.2. Distribuciones Continuas

2.2.1. Distribución Uniforme o Rectangular

Se dice que una **v.a.** continua X tiene distribución uniforme en el intervalo $[a, b]$, si su función de densidad está dada por:

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & a \leq x \leq b \\ 0; & \text{en otro caso.} \end{cases}$$

Parámetros: a, b

La gráfica de esta función se muestra en la siguiente figura: La **F.D.A.** de X, distribuida uniformemente, está dada por:

$$F(x) = P(X \leq x) = \begin{cases} 0; & x < a \\ \displaystyle\int_{-\infty}^{x} \frac{1}{b-a} dx = \frac{x-a}{b-a}; & a \leq x < b \\ 1; & x \geq b \end{cases}$$

La gráfica de F es:

$$X \sim U(x; a, b) \Rightarrow \begin{cases} E(x) = \mu = \dfrac{a+b}{2} \\ V(x) = \sigma^2 = \dfrac{(b-a)^2}{12} \\ D(x) = \sigma = \dfrac{(b-a)\sqrt{3}}{6} \end{cases}$$

2.2.2. Distribución Exponencial

Sea X una **v.a.** continua. Se dice que X tiene distribución exponencial con el parámetro real λ, si su función de densidad está dada por:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, & x \geq 0 \\ 0, & \text{en otro caso} \end{cases}$$

Parámetro: $\lambda > 0$

La gráfica de f, se muestra en la siguiente figura:

La **F.D.A.** de X, distribuida exponencialmente está dada por:

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ \displaystyle\int_{-\infty}^{x} \lambda \cdot e^{-\lambda \cdot x} dx = 1 - e^{-\lambda \cdot x}, & x \geq 0 \end{cases}$$

La gráfica de F, es la siguiente:

$$X \sim exp(x; \lambda) \Rightarrow \begin{cases} E(x) = \mu = \frac{1}{\lambda} \\ V(x) = \sigma^2 = \frac{1}{\lambda^2} \\ D(x) = \sigma = \frac{1}{\lambda} \end{cases}$$

Propiedad Amnésica

$$P(X > s + t/x > s) = P(X > t)$$

2.2.3. Distribución Normal

Una **v.a.** continua X tiene distribución normal con media $E(x) = \mu \in \mathbb{R}$ y varianza $V(x) = \sigma^2 > 0$, si su función de densidad está dada por:

$$f(x) = N(x; \mu, \sigma^2) = N(\mu, \sigma^2) = \frac{1}{\sigma \cdot \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})}, x \in \mathbb{R}$$

La gráfica de $f(x)$ es la siguiente:

Características de la Curva Normal

1. $f'(x) = 0 \Leftrightarrow x = \mu$
2. $x < \mu \Rightarrow f'(x) > 0$; luego f es creciente en $]-\infty, \mu]$
3. $x > \mu \Rightarrow f'(x) < 0$; luego f es decreciente en $[\mu, +\infty[$
4. La curva tiene un máximo en $x = \mu$. (Moda en μ)
5. $f''(x) = 0 \Leftrightarrow x = \mu + \sigma$
6. Los puntos de inflexión están en: $x = \mu - \sigma$ y $x = \mu + \sigma$.
7. La curva de la distribución normal es simétrica respecto de μ.
8. La media, mediana y moda son iguales (coinciden).
9. El área total bajo la curva normal y arriba del eje horizontal, es igual a 1.
10. La curva de la distribución normal se extiende $]-\infty, +\infty[$.

Distribución Normal Estándar

2.2.4. Distribución Chi Cuadrado

2.2.5. Distribución T-Student

2.2.6. Distribución F (Fisher-Snedecor)

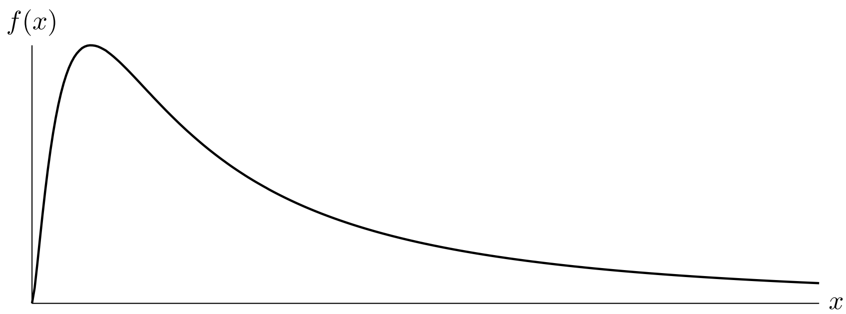
Una v.a. contínua X tiene distribución F, con r_1 grados de libertad en el numerador y r_2 grados de libertad en el denominador y se representa por:

$$X \sim F(r_1, r_2)$$

Si su función de densidad está definida por:

$$f(z) = \begin{cases} \dfrac{\left(\dfrac{r_1}{r_2}\right)^{\frac{r_1}{2}} \Gamma\left(\dfrac{r_1 + r_2}{2}\right)}{\Gamma\left(\dfrac{r_1}{2}\right)\Gamma\left(\dfrac{r_2}{2}\right)} \cdot \dfrac{x^{\frac{r_1}{2}-1}}{\left(1 + \dfrac{r_1}{r_2}x\right)^{\frac{r_1+r_2}{2}}} & ; x > 0 \\ 0 & ; \text{otro caso} \end{cases}$$

La siguiente figura muestra una curva particular de F, para valores particulares de r_1, r_2:



Teorema

Si U y V son v.a. independientes, tales que $U \sim \chi^2(r_1)$ y $V \sim \chi^2(r_2)$ entonces:

$$X = \frac{\frac{U}{r_1}}{\frac{V}{r_2}} \sim F(r_1, r_2)$$

Corolario

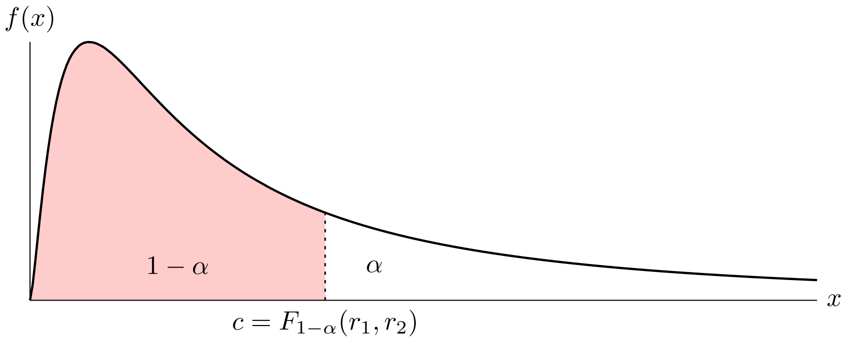
$$X \sim F(r_1, r_2) \Rightarrow \begin{cases} E(x) = \mu = \dfrac{r_2}{r_2 - 2} & \forall r_2 > 2 \\ \\ V(x) = \sigma^2 = \dfrac{2r^2(r_1 + r_2 - 2)}{r_2(r_2 - 2)^2(r_2 - 4)} & \forall r_2 > 4 \end{cases}$$

★ **Manejo de la Tabla:**

Si $X \sim F(r_1, r_2)$, entonces en la tabla de probabilidad F, se puede encontrar una probabilidad $1 - \alpha$ o un valor $c = F_{1-\alpha}(r_1, r_2)$ mediante la relación:

$$P(X \leq c) = 1 - \alpha \qquad \vee \qquad P(X \leq F_{1-\alpha}(r_1, r_2)) = 1 - \alpha$$

como se indica en la siguiente figura:



Para determinar los valores de F, correspondientes a areas:

$$1 - \alpha = 0{,}005, 0{,}010, 0{,}025, 0{,}05$$

o para determinar probabilidades correspondientes a valores de $c < 1$ se usa el teorema siguiente.

Teorema

Si:

$$X \sim F(r_1, r_2) \text{ entonces } \frac{1}{X} \sim F(r_2, r_1)$$

esto es:

$$F_{1-\alpha}(r_1, r_2) = \frac{1}{F_{\alpha}(r_2, r_1)}$$

2.2.7. Distribución Gamma

2.2.8. Distribución Beta

2.2.9. Distribución de Weibull

Capítulo 3

Muestreo y Distribuciones de Muestreo

3.1. Distribuciones de Muestreo

3.1.1. Estadígrafos y Estadísticos

Estadígrafos

Es todo número $\widehat{\Theta}$ obtenido a partir de los datos muestrales con el proposito de estimar los parámetros poblacionales.

$$\widehat{\Theta}_1 = f(x_1, x_2, \ldots, x_n)$$

donde $x_1, x_2, \ldots, x_n$ son los valores muestrales de las correspondientes v.a. $X_1, X_2, \ldots, X_n$, es decir un estadigrafo es un número obtenido por una función únicamente de las v.a.

$$X_1, X_2, \ldots, X_n$$

Estadística o Estadístico

Si seleccionamos k muestras aleatorias de la misma población, obtendremos k valores para $\widehat{\Theta}$, tambien aleatorios. Así $\widehat{\Theta}$ es a su vez una v.a. llamada estadística o estadístico cuyos valores son estadígrafos.

3.1.2. Desigualdad Chebyshev

Si X es una v.a. con una media μ y varianza σ^2 entonces:

$$P(|x-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}; \qquad \text{donde } k > 0$$

Si X es v.a. normal, entonces:

$$\begin{aligned} P(|x-\mu| < k\sigma) =& P(-k\sigma < x-\mu < k\sigma) \\ =& P(-k < z < k) \\ =& \emptyset(k) - \emptyset(-k) \\ =& \emptyset(k) - (1 - \emptyset(k)) \\ =& 2\emptyset(k) - 1 \end{aligned}$$

$$\therefore P(|x-\mu| < k\sigma) = 2\emptyset(k) - 1$$

3.2. Distribución de Muestreo

Si $\widehat{\Theta} = \overline{x}$ la distribución de la estadística $\overline{x}$ se llama **distribución de la media muestral** o distribución de muestreo de la media. Si $\widehat{\Theta} = S^2$, la distribución de la estadística S^2 se llama **distribución de la varianza muestral** o distribución de muestreo de la varianza.

3.3. Distribución de la Media Muestral

Si:

$$X \sim f(x;\mu,\sigma^2)$$

donde: $E(x)=\mu$ y $V(x)=\sigma^2$, entonces:

$$E(\overline{x})=\mu_{\overline{x}}=\mu$$

$$V(\overline{x})=\sigma^2_{\overline{x}}=\begin{cases}\dfrac{\sigma^2}{n}; & \text{si la poblacion es infinita}\\ \\ \dfrac{\sigma^2}{n}\left(\dfrac{N-n}{N-1}\right); & \text{si la poblacion es finita de tamaño } N.\end{cases}$$

donde: $\dfrac{N-n}{N-1}$ es el factor de corrección por poblacion finita.

Demostración

Sea $\{x_1,x_2,\ldots,x_n\}$ donde $x_1,x_2,\ldots,x_n$ son los valores muestrales de las v.a.'s independientes tales que:

$$\forall i\in I\begin{cases}E(x_i)= & \mu\\ V(x_i)= & \sigma^2\end{cases}$$

$$\begin{aligned}E(\overline{x})=&E\left(\frac{\sum_{i=1}^{n}x_i}{n}\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}x_i\right)\\ =&\frac{1}{n}\left(\sum_{i=1}^{n}E(x_i)\right)\\ =&\frac{1}{n}\sum_{i=1}^{n}\mu\\ =&\frac{1}{n}(n\mu)\\ =&\mu\end{aligned}$$

$\therefore E(\overline{x})=\mu$

3.3.1. Distribución de Muestras de la diferencia o suma de Medias

(I) Si $\overline{x},\overline{y}$ son las medias de m.a.'s independientes de tamaño n_1,n_2, extraidas de dos poblaciones infinitas con medias μ_x,μ_y y varianza σ^2_x,σ^2_y, respectivamente:

$$z=\frac{(\bar{x}\pm\bar{y})-(\mu_x\pm\mu_y)}{\sqrt{\dfrac{\sigma^2_x}{n_1}+\dfrac{\sigma^2_y}{n_2}}}$$

es una v.a cuya distribución se **aproxima** a la normal estándar cuando $n_1,n_2\to\infty$.
Si las poblaciones son normales z es una v.a. cuya distribución es **exactamente** la normal estándar, aun cuando n_1 o n_2 sean pequeños.
Si las poblaciones son finitas de tamaño N_1,N_2, respectivamente, entonces:

$$z=\frac{(\bar{x}\pm\bar{y})-(\mu_x\pm\mu_y)}{\sqrt{\dfrac{\sigma^2_x}{n_1}\left(\dfrac{N_1-n_1}{N_1-1}\right)+\dfrac{\sigma^2_y}{n_2}\left(\dfrac{N_2-n_2}{N_2-1}\right)}}$$

(II) Distribución t-Student

Distribución de la diferencia o suma de medias, cuando las varianzas poblacionales son desconodidas, muestras pequeñas y poblaciones normales homoceclásticas.
Si $\overline{x}, \overline{y}$, son las medias de m.a.'s independientes de tamaño n_1, n_2 (ambos pequeños), extraídos de dos poblaciones normales homoceclasticas, con media μ_x, μ_y y varianzas poblacionales desconocidas, entonces:

$$t = \frac{(\bar{x} \pm \bar{y}) - (\mu_x \pm \mu_y)}{\sqrt{\dfrac{(n_1 - 1)\hat{S}_x^2 + (n_2 - 1)\hat{S}_y^2}{n_1 + n_2 - 2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

es el valor de una v.a. cuya distribución es la t-Student con: $r = n_1 + n_2 - 2$ **g.l.**

(III) Si las muestras grandes y las varianzas muestrales son desconocidas estas se pueden estimar a partir de las varianzas muestrales ordinarias:

$$\sigma_x^2 \simeq S_x^2, \sigma_y^2 \simeq S_y^2$$

Y en lugar de usar t, se usa z, y esto es valido aun en los casos en los que las poblaciones no sean normales.

3.3.2. Distribución de la Proporción Muestral

(I) Si la estadística $\widehat{p}$ es la proporción de éxitos en una m.a. de tamaño n extraída de una población binomial infinita con proporción de éxitos p, entonces:

$$z = \frac{\widehat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

es una v.a. cuya distribución se aproxima a la normal estándar cuando $n \to \infty$.
Si la población binomial es finita de tamaño N, entonces:

$$z = \frac{\widehat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}\left(\dfrac{N-n}{N-1}\right)}$$

Ademas, como se pasa de una v.a. discreta (binomial) a una continua (normal), se debe introducir el factor de corrección de cantidad $\dfrac{1}{2n}$, sumando este factor al límite superior y restando al límite inferior.

$$\hat{p} = \frac{x}{n}$$

Si n es suficientemente grande, se puede obviar el factor de corrección:

$$E(\hat{p}) = \mu_{\hat{p}} = p$$

$$V(\hat{p}) = \sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$$

(II) Si las estadísticas $\widehat{p_1}, \widehat{p_2}$ son las proporciones de éxito en m.a.'s n_1, n_2 extraídas de dos poblaciones binomiales infinitas con proporción de éxito p_1, p_2 respectivamente, entonces:

$$z = \frac{(\hat{p_1} - \hat{p_2}) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}}$$

es una v.a. cuya distribución se aproxima a la normal estandar cuando $n_1, n_2 \to \infty$.
Si las poblaciones binomiales son finitas de tamaño N_1, N_2 respectivamente, entonces:

$$z = \frac{(\hat{p_1} - \hat{p_2}) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1}\left(\dfrac{N_1 - n_1}{N_1 - 1}\right) + \dfrac{p_2(1-p_2)}{n_2}\left(\dfrac{N_2 - n_2}{N_2 - 1}\right)}}$$

Capítulo 4

Estimación Estadística

La Estimación Estadística es parte de la inferencia estadistica, cuyo objetivo es estimar los parámetros de una población mediante muestras aleatorias provenientes de ellas. Existen básicamente dos formas de estimación de un parámetro poblacional: Estimación Puntual y Estimación por Intervalo.

4.1. Estimación Puntual

Sean $X_1, X_2, \ldots, X_n$ m.a.'s de tamaño n, seleccionada de una población, cuya distribución es $f(x, \theta)$ siendo θ el parámetro. Se denomina **Estimador Puntual** del parámetro θ a cualquier estadístico (v.a.) $\widehat{\theta} = H(X_1, X_2, \ldots, X_n)$, cuyo valor numérico (estadígrafo) proporcionará una **Estimación Puntual** del parametro θ.

Parámetros	Estadígrafos
μ	$\bar{x}$
σ^2	S^2 ó $\hat{S}^2$
p, π	$\hat{p}, \hat{\pi}$

4.1.1. Propiedades de los Estimadores Puntuales

No toda función de la muestra es un buen estimador de parametro. Un buen estimador es aquel que está mas cerca del parametro que se estima. Para que un estimador puntual sea bueno, debe poseer ciertas propiedades como:

- Insesgadez (Insesgabilidad)
- Eficiencia
- Consistencia
- Suficiencia

Estimador Insesgado

$\widehat{\Theta}$ es un estimador Insesgado del parametro Θ:

$$E(\widehat{\Theta}) = \Theta$$

Estimador Eficiente

$\widehat{\Theta}_1$ es un estimador mas eficiente que $\widehat{\Theta}_2$ (ambos insesgados) del parametro Θ ssi:

$$V(\widehat{\Theta}_1) < V(\widehat{\Theta}_2)$$

Estimador Consistente

$\widehat{\Theta}$ es un estimador consistente del parametro Θ ssi:

$$\begin{cases} \textbf{i)} \lim\limits_{x\to\infty} E(\widehat{\Theta}) = 0 \\ \textbf{ii)} \lim\limits_{x\to\infty} V(\widehat{\Theta}) = 0 \end{cases}$$

Estimador Suficiente

$\widehat{\Theta}$ es un estimador suficiente con parametro Θ ssi:

"$\widehat{\Theta}$ aporta tanta informacion acerca de los parametros que se estiman, tomando una muestra del parametro."

4.1.2. Sesgo y Error Cuadrático Medio de un Estimador (ECM)

Sesgo

Sesgo del estimador $\widehat{\Theta}$:

$$Sesgo(\widehat{\Theta}) = E(\widehat{\Theta}) - \Theta$$

Si $Sesgo(\widehat{\Theta}) = 0$, entonces $\widehat{\Theta}$ es un estimador insesgado:

EMC

$$ECM(\Theta) = E[(\widehat{\Theta} - \Theta)^2] = V(\Theta) + (Sesgo(\Theta))^2$$

4.2. Estimación por Intervalo

Consiste en la estimación de un parámetro poblacional mediante la obtención de un intervalo aleatorio llamado **Intervalo de Confianza** cuyos limites superior e inferior:

$$L_i \wedge L_s$$

Son funciones de las v.a. observadas que dependen de cierto estimador $\widehat{\Theta}$ parametro Θ con una probabilidad:

$$1 - \alpha$$

que representa el nivel o coeficiente de confianza deseado, esto es, probabilidad de que el parámetro esté comprendido:

$$\mathcal{P}(L_i \leq \Theta \leq L_s) = 1 - \alpha$$

El objetivo de la estimación por intervalo es obtener, el intervalo:

$$L_i \leq \Theta \leq L_s$$

y se espera que $100(1-\alpha)\,\%$ de ello, abarque el parametro de la población Θ.

4.2.1. Formulación General de los Intervalos

Consideremos $z = \dfrac{\widehat{\Theta} - \Theta}{\sigma_\Theta}$ de aquí; se tiene:

$$\mathcal{P}\left(-z_{1-\frac{\sigma}{2}} \leq z \leq z_{1-\frac{\sigma}{2}}\right) = 1 - \sigma$$

$$\Leftrightarrow \mathcal{P}\left(-z_{1-\frac{\sigma}{2}} \leq \frac{\widehat{\Theta} - \Theta}{\sigma_\Theta} \leq z_{1-\frac{\sigma}{2}}\right) = 1 - \sigma$$

$$\Leftrightarrow \mathcal{P}\left(-z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta \leq \widehat{\Theta} - \Theta \leq z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta\right) = 1 - \sigma$$

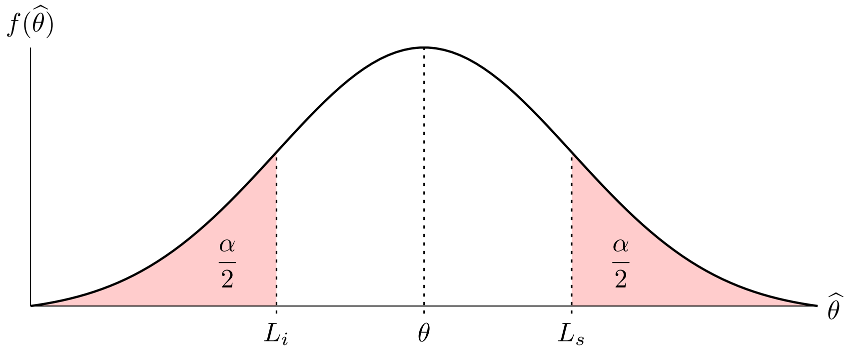
$$\Leftrightarrow \mathcal{P}\left(-z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta - \widehat{\Theta} \leq -\Theta \leq z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta - \widehat{\Theta}\right) = 1 - \sigma$$

$$\Leftrightarrow \mathcal{P}\left(z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta + \widehat{\Theta} \geq \Theta \geq -z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta + \widehat{\Theta}\right) = 1 - \sigma$$

$$\Leftrightarrow \mathcal{P}\left(-z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta + \widehat{\Theta} \leq \Theta \leq z_{1-\frac{\sigma}{2}} \cdot \sigma_\Theta + \widehat{\Theta}\right) = 1 - \sigma$$

4.2.2. Interpretación de Intervalo de Confianza

Si se selecciona repetidamente 100 veces muestras de tamaño n tendremos 100 intervalos semejantes al estimador $\widehat{\Theta} \pm z_{\pm\frac{\alpha}{2}} \cdot \sigma_{\widehat{\Theta}}$ y esperamos que $100(1-\alpha)\,\%$ incluyan al parametro y $100\alpha\,\%$ no lo incluirá.



4.2.3. Estimación por Intervalo de la media Poblacional

(I) $\mathcal{P}\left(\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$

La estadística z se usa en los siguientes casos:

- **(i)** Muestra grande, Varianza Poblacional conocida y Poblacion normal o no.
- **(ii)** Muestra pequeña, Varianza Poblacional conocida y población normal.
- **(iii)** Muestra grande, Varianza desconocida y poblacion normal o no.

(II) $\mathcal{P}\left(\bar{x} - t_{1-\frac{\alpha}{2}}(r) \cdot \frac{\widehat{S}}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{1-\frac{\alpha}{2}}(r) \cdot \frac{\widehat{S}}{\sqrt{n}}\right) = 1-\alpha$

La estadística t se usa cuando:

- **(i)**
- **(ii)** La varianza poblacional es desconocida.
- **(iii)** La poblacion debe ser normal

Donde: $r = n-1$ **g.l.**

4.2.4. Estimación por Intervalo de la Diferencia o suma de medias

(I) $\mathcal{P}\left((\bar{x} \pm \bar{y}) - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}} \leq \mu_x \pm \mu_y \leq (\bar{x} \pm \bar{y}) + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}\right) = 1-\alpha$

La Estadística z se aplica en los siguientes casos:

- **(i)** Muestras grandes, varianza poblacional conocida y población normal o no.
- **(ii)** Muestra pequeña, varianzas poblacionales conocidos y población normal.
- **(iii)** Muestra grande, variables poblacionales desconocidos y población normal o no.

$$\sigma_x^2 \simeq S_x^2 \quad \wedge \quad \sigma_y^2 \simeq S_y^2$$

(II) $\mathcal{P}\left((\bar{x} \pm \bar{y}) - t_{1-\frac{\alpha}{2}}(r) \cdot \beta \leq \mu_x \pm \mu_y \leq (\bar{x} \pm \bar{y}) + t_{1-\frac{\alpha}{2}}(r) \cdot \beta\right) = 1-\alpha$

La Estadística t se usa cuando:

- **(i)** La muestra pequeña.

(ii) Varianzas poblacionales desconocidas.

(iii) Poblaciones son normales.

Donde:

- $\beta = \sqrt{\frac{(n_1-1)\hat{S_x}^2 + (n_2-1)\hat{S_y}^2}{n_1+n_2-2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$
- $r = n_1 + n_2 - 2$ **g.l.**

4.2.5. Estimación por Intervalo para Proporciones

(I) $\mathcal{P}\left(\hat{p} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1-\alpha$

(II) $\mathcal{P}\left((\widehat{p_1} \pm \widehat{p_2}) - z_{1-\frac{\alpha}{2}} \cdot \beta \leq p_1 \pm p_2 \leq (\widehat{p_1} \pm \widehat{p_2}) + z_{1-\frac{\alpha}{2}} \cdot \beta\right) = 1-\alpha$

Donde:

- $\beta = \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$

Si alguna de las poblaciones binomiales son finitos se debe incorporar el factor de correción por población finita.

Tabla: Coeficientes de Confianza o valores críticos $(1 - \frac{\alpha}{2})$

$1-\alpha$	0.9973	0.99	0.98	0.96	0.9545	0.95	0.90	0.68
$100(1-\alpha)\,\%$	99.73	99	98	96	95.45	95	90	68
$z_{1-\frac{\alpha}{2}}$	3.00	2.58	2.33	2.05	2.00	1.96	1.64	1.00

4.2.6. Relación entre el Error de Estimación y el Riesgo de Estimación

$$E = |\widehat{\Theta} - \Theta| \text{ (Estimador - Parámetro)}$$

$$z = \frac{\widehat{\Theta} - \Theta}{\sigma_{\widehat{\Theta}}}$$

$$-z_{1-\frac{\alpha}{2}} \leq z \leq z_{1-\frac{\alpha}{2}}$$

$$\Updownarrow$$

$$-z_{1-\frac{\alpha}{2}} \leq \frac{\widehat{\theta} - \theta}{\sigma_{\widehat{\theta}}} \leq z_{1-\frac{\alpha}{2}}$$

$$-z_{1-\frac{\alpha}{2}} \cdot \sigma_{\widehat{\theta}} \leq \widehat{\theta} - \theta \leq z_{1-\frac{\alpha}{2}} \cdot \sigma_{\widehat{\theta}}$$

$$|\widehat{\theta} - \theta| < z_{1-\frac{\alpha}{2}} \cdot \sigma_{\widehat{\theta}}$$

$$E < z_{1-\frac{\alpha}{2}} \cdot \sigma_{\widehat{\theta}}$$

4.2.7. Relación entre el Error de Estimación y el tamaño de la Muestra

(I) En la Distribución de la **Media Muestral**

$$\begin{aligned} E = |\overline{x} - \mu| &\Rightarrow E \geq z_{1-\frac{\alpha}{2}} \sigma_{\overline{x}} \\ &\Rightarrow E \geq z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \\ &\Rightarrow \sqrt{n} \geq \frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{E} \\ &\Rightarrow n \geq \left(\frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{E} \right)^2 \end{aligned}$$

(II) En la Distribución de la **Proporción Muestral**

- $n = \dfrac{z^2_{1-\frac{\alpha}{2}} \cdot p \cdot q}{E^2}$

  Población Infinita

- $n = \dfrac{z^2_{1-\frac{\alpha}{2}}}{4E^2}$

  $q = p = 0{,}5$

- $n = \dfrac{N \cdot z^2_{1-\frac{\alpha}{2}} \cdot p \cdot q}{(N-1)E^2 + z^2_{1-\frac{\alpha}{2}} \cdot p \cdot q}$

  Población Finita