HOJA N°

FECHA 1

ESTADISTICA

- Frecuencia Absoluta — Suma total de datos
- Frecuencia Relativa — $\frac{\text{Frecuencia Absoluta}}{\text{Total de datos}}$
- Frecuencia Porcentual — Frecuencia Relativa × 100
- Frecuencia Acumulada — Frec pasada + Frec actual
- Ancho aprox de clase — $\frac{\text{Dato mayor} - \text{Dato menor}}{\text{Num clases}}$

Graficos.

Tallo y Hoja
- acomodar el primer digito a la izquierda verticalmente
- a la derecha se anota el ultimo digito

1	3 2 4
2	5 6
3	7 7 6

Caja y bigote
- determinar el valor max y min
- se determina en el Q3 y Q1
- el Q2 es la mediana

Estadistica descriptiva.

- Mediana: valor central de los datos ordenados
- Media (promedio): $\frac{\text{suma de los valores}}{\text{Q valores}}$

$$\text{COEF. V} = \frac{\text{Desvest}}{\text{Media}} \times 100$$

- Moda: valor con mas frecuencia.
- Percentil: $\left(\frac{P}{100}\right) \times n$ = posición

$Q_1 = P_{25}$
$Q_2 = P_{50}$ (mediana)
$Q_3 = P_{75}$

Medidas de varianza

- Amplitud total: Max − Min

Esperanza (promedio): $E(x)\ X_1 \cdot p_1 + X_2 \cdot p_2$.

$E(x^2) - E^2$

- Rango intercuartil: $Q_3 - Q_1$
- Varianza: $\frac{\text{suma de datos}}{\text{Q datos}}$; $\frac{(\text{suma de muestra}^2) - Q \cdot E^2}{\text{Q datos} - 1}$
- Desvio estandar: $\sqrt{\text{varianza}}$

Combinatoria $C^N_R = \binom{N}{R} = \frac{N!}{R!\,(N-R)!}$ (1 2 3)

$P^N_R = \binom{N}{R}\ \frac{N!}{(N-R)!}$ (TODAS)

NOTA

PROBABILIDAD

$X \sim H(N; M; n)$

- **HIPERGEOMETRICA**: TOTALES ASIGNADOS. POBLACION FINITA.

$$p(r) = \frac{\binom{M}{r}\binom{N-M}{n-r}}{\binom{N}{n}}$$

- $E(x) = n \times \frac{M}{N}$
- $V = \left(\frac{N-n}{n-1}\right) n \times p \times (1-p)$

SUCESO ALEATORIO

- → OPUESTO A'
- → CONJUNTO $A \cap B$
- → EXCLUYENTE $A \cup B$
- → A PERO NO B $A \cap B'$
- → UNO DE AMBOS $(A \cap B') \cup (A' \cap B)$
- → TODO MENOS INTERSECCION $A' \cap B'$
- → TODO lo DE AFUERA $A' \cup B'$

- **POISSON** μ: EXITOS EN UN PERIODOS

$$p(r) = \frac{\mu^r \cdot e^{-\mu}}{r!}$$

$\mu = \lambda$ ↳ tasa $\times$ t

$E = \mu$ $\quad V = \mu$

$X \sim P(\lambda; t)$

o • ADICION $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

y • CONDICIONAL $P(A \cap B) = P(A/B) \times P(B) \Rightarrow P(A/B) = \frac{P(A \cap B)}{P(B)}$

si • BAYES $P(B/A) = \frac{P(A/B) \times P(B)}{P(A)}$

- TOTAL $P(A) = P(A/B) \times P(B) + P(A/C) \times P(C)$
- EVENTOS INDEPENDIENTES $P(A \cap B) = P(A) \times P(B)$

VARIABLES ALEATORIAS DISCRETAS

DISCRETA se le asigna un numero a cada elemento del recorrido

- **BERNOULI**: 1 repeticion, 2 posibles resultados $\begin{cases} 1 & \text{exito} \\ 0 & \text{fracaso} \end{cases}$

X = Fue exitoso o No $\quad X \sim B(1; P)$

p = Probabilidad de exito $\quad R(x) = \{0; 1\}$ $\quad V = P \cdot (1-P)$

$$P(x=1) = p \qquad P(x=0) = 1-p$$

- **BINOMIAL**: m repeticiones, 2 posibles resultados (POBLACIÓN ∞)

X = cant [illegible] en m repeticiones $\quad X \sim Bi(m; p)$

p = Probabilidad de exito

$$P_X(r) = \binom{m}{r} p^r (1-p)^{m-r}$$

r	0
$\binom{n}{r}$	$\binom{n}{0}$

- ESPERANZA $m \times p + m \times p + m \times p \ldots$
- VARIANZA $m \times p \times (1-p)$
- DESVIO EST $\sqrt{m \times p \times (1-p)}$

N = POBLACION TOTAL
n = MUESTRA
M = EXITOS en la poblacion
p = probabilidad exito
r = Valor aleatorio
λ = Tasa $\times$ UNIDADES.

Scanned with CamScanner

HOJA N°

FECHA ②

Variables Aleatorias Continuas

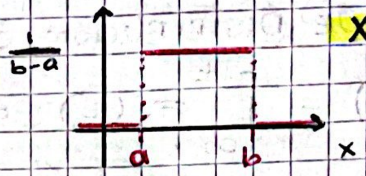
- **UNIFORME**: una probabilidad es proporcional a la longitud del intervalo

Función de densidad poblacional

(Buscar el area con la Integral)

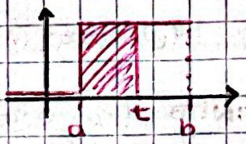
$X \sim U(a;b)$

$$F(x)\begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otros casos} \end{cases}$$



Función de distribución o acumulada

$$p(x \leq t) = \frac{t-a}{b-a}$$



- base $= t-a$
- altura $= \dfrac{1}{b-a}$
- AREA $= b \times h = (t-a) \cdot \dfrac{1}{b-a}$

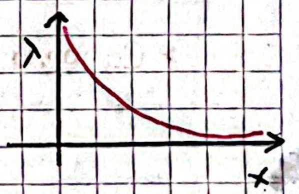
$$E(x) = \frac{a+b}{2} \qquad VAR(x) = \frac{(b-a)^2}{12}$$

- **EXPONENCIAL**: probabilidad acerca del tiempo $X \sim E(\lambda)$

Función de densidad poblacional

$\dfrac{1}{\mu} = \lambda$

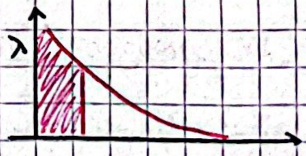
$$F(x)\begin{cases} \dfrac{1}{\mu} \cdot e^{-\frac{x}{\mu}} & x \geq 0 \ ; \ \mu > 0 \\ 0 & \text{otros casos} \end{cases}$$



Función de distribución o acumulada

$$p(x \leq t) = 1 - e^{-\lambda \cdot t}$$

$$p(x > t) = e^{-\lambda \cdot t}$$



$$E(x) = \frac{1}{\lambda} = \mu \qquad VAR(x) = \frac{1}{\lambda^2} = \mu^2$$

POISSON

NOTA

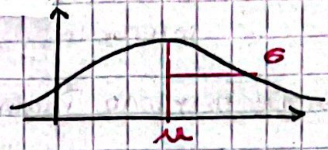
• **NORMAL**: tiene una media de cero y una desviación standar de uno

MEDIA POBLACIONAL.

$\mu = 0$ y $\sigma = 1$ $\qquad X \sim N(0;1)$ $(\mu;\sigma)$

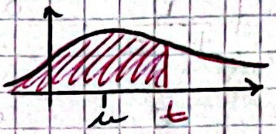
FUNCIÓN DE DENSIDAD NORMAL

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$



FUNCIÓN DE DISTRIBUCIÓN O ACUMULADA

$$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} F_X(t)\, dx \longrightarrow \text{Probability}$$



$\mu = n.p \qquad \sigma = \sqrt{n.p.q} \qquad Z = \frac{X-\mu}{\sigma}$ → TAMAÑO del EFECTO.

MUESTREO Y DISTRIBUCIÓN MUESTRAL

Hay población de tipo Finita (se aclara) e Infinita (no aclara)

error estandar

• **Población Finita**: se toma una muestra (n) de una población finita de tamaño (N) donde cada posible muestra tiene las mismas posibilidades

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \left(\frac{\sigma}{\sqrt{n}}\right)$$

• **Población Infinita**: se selecciona una muestra (n) de la población y cada elemento es Independiente

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

FORMAS DE LA DISTRIBUCIÓN MUESTRAL

$\bar{X}$ → media Muestral.

> Conozco la varianza: $X \sim N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$

$$X = Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \qquad Z \sim N(0,1)$$

VALOR - P.

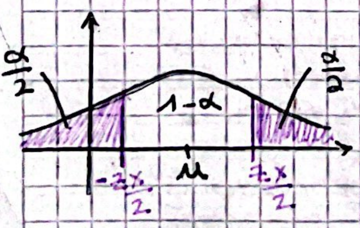
valor crítico.

> Desconozco la varianza: $Z = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ $\quad (n \geq 30)$ uso Normal.

S = desvío desconocido

↳ basado por los datos.

T de student ⟵ $t = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ $\qquad t \sim t(n-1)$ $\quad (n < 30)$



$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\frac{\alpha}{2}}\right)$$

$$P\left(\bar{X} - \boxed{Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

ERROR DE LA MEDIA

INTERVALOS DE CONFIANZA:

NOTA

se tomo un n aleatorio y uso

HOJA N°

FECHA ③

Error Estandar

DESVIACION ESTANDAR de la proporcion Muestral $\hat{p}$

INT. CONF $(\hat{p} \pm _{1-\frac{\alpha}{2}} \cdot \text{error est})$

- POBLACION FINITA: $\sigma_{\hat{p}} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{p(1-p)}{n}}$
- POBLACIÓN INFINITA: $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

$p \sim N\left(p ; \sqrt{\frac{p(1-p)}{n}}\right)$

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

$\text{VAR}(\hat{p}) = \sigma^2_{\hat{p}} = \frac{p(1-p)}{n}$

$\text{DESV}(\hat{p}) = \sigma_{\hat{p}}$ = ERROR ESTANDAR DE LA PROPORCION

$Z = N(0;1) \sim \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ → (ESTADISTICO DE PRUEBA)

Z: VALOR CRITICO

PROPIEDADES DE LOS ESTIMADORES

- INSESGADEZ: Valor esperado muestral es igual a parametro de la poblacion que se estudia
  $E(\hat{\theta}) = \theta$
- EFICIENCIA: estimador puntual con menor error estandar
- CONSISTENCIA: el valor estimado tiende a estar mas cerca del parametro poblacional a medida que el tamaño de la muestra aumenta.

ESTIMACION POR INTERVALOS → CONFIANZA $= \bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

NIVEL DE CONFIANZA $= P(\bar{X} - \text{error estimacion} < \mu < \bar{X} + \text{error estimacion})$

ERROR ESTIMACION σ CONOCIDA $= Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

ERROR ESTIMACION σ DESCONOCIDA $= t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$

INT. C = nivel de conf (%)
COEF. C = INT. C / 100 → α → $\frac{\alpha}{2}$
$Z_{\frac{\alpha}{2}}$ = valor a la app (σ)

SACAR Z
1. INTERVALOR. C / 100 = α
2. COEF. C / 2 = $\alpha/2$
3. $1 - \alpha/2$
4. El COEF α es el que se pone en $P(X > x)$ en la app

TAMAÑO DE LA MUESTRA PARA UNA ESTIMACION DE LA MEDIA POBLACIONAL
$$n = \frac{\left(Z_{\frac{\alpha}{2}}\right)^2 \cdot \sigma^2}{E^2}$$
E → margen de ERROR

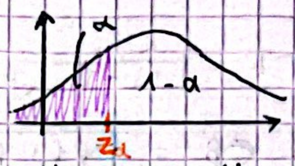
TAMAÑO DE LA MUESTRA PARA UNA ESTIMACION DE LA PROPORCION POBLACIONAL
$$n = \frac{\left(Z_{\frac{\alpha}{2}}\right)^2 \cdot p^{\circ} \cdot (1 - p^{\circ})}{E^2}$$
p° → valor $\hat{p}$

NOTA

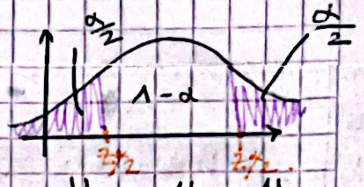
Prueba de Hipotesis

- Afirmacion referida a una informacion POBLACIONAL que proporciona un criterio para tomar diciciones.

$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$



$H_0: \mu \geq \mu_0$
$H_1: \mu < \mu_0$



$H_0: \mu = \mu_0$
$H_1: \mu \neq \mu_0$

Regla de Decisión

$z_\alpha \rightarrow$ % confianza

(I)
$$\begin{cases} \bar{x} \geq \mu + z_\alpha \cdot \dfrac{\sigma}{\sqrt{n}} & \text{Rechaza } H_0 \\ \bar{x} < \mu + z_\alpha \cdot \dfrac{\sigma}{\sqrt{n}} & \text{No rechaza } H_0 \end{cases}$$

Errores de tipo I y II

	Ho verdadero	Ho Falso
Rechaza Ho	Error tipo I α	$1-\beta$
No rechaza Ho	$1-\alpha$	Error tipo II β

Error tipo I → Rechaza Ho siendo verdadero

Error tipo II β → No Rechaza Ho siendo Falso.

α = Nivel significancia (probabilidad de cometer error I)

- Valor P = minimo α que Rechaza a Ho en base a los datos de la muestra

(determina si Ho se rechaza)

$V_p \leq \alpha$ – Rechaza Ho
$V_p > \alpha$ – No Rechaza Ho

→ Encontrar Z → $Z \sim N(0,1)$ // $\bar{x} \sim N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$

Tamaño del Efecto: $\dfrac{|\bar{x} - \mu|}{\sigma}$ $\qquad \dfrac{|\hat{p} - p|}{\sqrt{\hat{p}(1-\hat{p})}}$

- Valor Critico = Valor z que corresponde al α

↳ $\alpha \Rightarrow z = (z_\alpha$ / $z_{\alpha/2}$ Tabla de prob estandar)

$\boxed{\text{Error tipo II}}$ (β) (pag 372 libro)

" Encontrar Z.

$$P\left(\bar{x} > \mu_0 \pm z_\alpha \cdot \frac{\sigma}{\sqrt{n}} \;\middle|\; \mu = 0\right)$$

$\bar{x} \sim N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$ $\qquad x > \mu_0 \pm z_\alpha \cdot \dfrac{\sigma}{\sqrt{n}}$.

NOTA:

Tabla Prob. Estandar.

1 cola:

α	z_α
0,01	2,3263
0,05	1,64485
0,1	1,281552

2 colas:

α	$z_{\alpha/2}$
0,01	2,575829
0,05	1,959964
0,1	1,644854

HOJA Nº

FECHA 4

se tomo un n aleatorio y usan metodo 1. se tomo n aleatorio y usan metodo 2

DIFERENCIAS DE MEDIAS PARA MUESTRAS INDEPENDIENTES.

		Dif
3	2	1
5	6	-1
2	4	-2

$\mu_1 = 3,\hat{3}$ $\quad \mu_2 = 4$ $\quad \mu = -0,\hat{6}$

$$\mu_1 - \mu_2 = 0,33 - 4 = 0,\hat{6}$$

- **Con Varianza Conocida:**

Estimador puntual entre medias: $\bar{X}_1 - \bar{X}_2$

Error Estandar: $\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Estimacion por Intervalos: $\bar{X}_1 - \bar{X}_2 \pm \underbrace{\dfrac{Z\alpha}{2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}_{\text{margen de error}}$

- Prueba de Hipotesis $\mu_1 - \mu_2$

$H_0: \mu_1 - \mu_2 \geq D_0 \qquad \mu_1 = \mu_2 \leftrightarrow \mu_1 - \mu_2 = 0.$

$H_1: \mu_1 - \mu_2 < D_0$

$$t \,|\, Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \quad \text{(Estadistico)}$$

$\sigma = s$

Tamaño del efecto: $g = \dfrac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\dfrac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}}}$

- **Con Varianza Desconocida:**

Estimacion por Intervalos: $\bar{X}_1 - \bar{X}_2 \pm t_{\frac{\alpha}{2}} \sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}$

Grado de libertad:

$$gl = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{S_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{S_2^2}{n_2}\right)^2}$$

NOTA

DIFERENCIA ENTRE PROPORCIONES POBLACIONALES.

- Error Estandar $\hat{p}_1 - \hat{p}_2$: $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$

- Estimacion por Intervalos : $\hat{p}_1 - \hat{p}_2 \pm \frac{Z\alpha}{2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
  $(1-\alpha)$ coef conf

- Error Estandar $P_1 = P_2 = P$: $\sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

- Estimador combinado : $\hat{P} = \frac{n_1 . \hat{P}_1 + n_2 . P_2}{n_1 + n_2}$ (Promedio ponderado de P_1 y P_2.)

- Estadistico : $Z = \frac{P_1 - P_2}{\sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

TAMAÑO DE UNA MUESTRA $n = \frac{(Z_\alpha + Z_\beta)^2 . \sigma^2}{(\mu_0 - \mu_a)^2}$.

DIFERENCIAS DE MEDIA PARA MUESTRAS PAREADAS.

Se toma un n aleatorio y usan ambos metodos con inicio aleatorio (empezar por 1 o 2)

MEDIA : $\frac{\Sigma(\overbrace{M_1 - M_2}^{\text{metodo } d_i})}{n} = \bar{d}$ PROMEDIO.

Desvio : $S_d = \sqrt{\frac{\Sigma(d_i - \bar{d})}{n-1}}$

Estadistico : $t = \frac{\bar{d} - \mu_0}{\frac{Sd}{\sqrt{n}}}$