FORMULARIO DE ESTADÍSTICA

Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América.

Facultad de Ciencias Económicas

Rufasto C. Andy

CONTENIDO

DISTRIBUCIÓN DE FRECUENCIAS

1.1 Distribuciónes de Frecuencias

R = $V_{max} - V_{min}$

C = $\dfrac{R}{K}$

K = $1 + 3.332 \log N$

$F = K * C \; - \; R$

R: Rango
V: Valor
C: Amplitud de intervalo
K: Número de clase
$k \quad \in \quad [5;20]$
F: Factor de centralización
N: Número de datos

1.2 Tabla de Frecuencias

	x_i	f_i	F_i	h_i	H_i
i=1	$[V_{min}; V_{min} + c\rangle$	f_1	f_1	f_1/N	F_1/N
i=2	$[L_{i2}; L_{i2} + c\rangle$	f_2	$F_1 + f_2$	f_2/N	F_2/N
i=3	$[L_{i3}; L_{i3} + c\rangle$	f_3	$F_2 + f_3$	f_3/N	F_3/N
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i=k	$[L_{ik}; L_{ik} + c\rangle$	F_k	N	f_k/N	1

Tabla 1.1: Tabla de Distribución de Frecuencias

MEDIDAS DE TENDENCIA CENTRAL

2.1 Media

2.1.1 Media Aritmética

Método Largo

$$\bar{x} = \sum_{i=1}^{N} \frac{x_i}{N}$$
$$= \sum_{i=1}^{N} \frac{x_i f_i}{N}$$

Método Corto

$$\begin{aligned} Si \quad x_i &= d_i + A \\ \bar{x} &= \overline{d + A} \\ &= \overline{d} + A \\ \bar{x}_i &= \sum_{i=1}^{N} \frac{x_i d_i}{N} + A \end{aligned}$$

Método Clave

$$\begin{aligned} d_i &= cu_i \\ x &= cu_i + A \\ \bar{x} &= c\left[\sum_{i=1}^{N} \frac{f_i u_i}{N}\right] + A \end{aligned}$$

2.1.2 Media Geométrica

$$G = \sqrt[n]{X_1^{f_1}.X_2^{f_2}.X_3^{f_3}\dots X_n^{f_n}}$$

2.1.3 Media Armónica

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i . f_i}}$$

2.2 Mediana

$$Med = L_i + c\left[\frac{\frac{n}{2} - F_{i-1}}{F_i - F_{i-1}}\right]$$

L_i: Límite inferior de la clase modal.
F_i: frecuencia acumulada del intervalo mediano.

2.3 Moda

$$\tilde{x} = L_i + c\left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right]$$

$$\Delta_1 : f_i - f_{i-1}$$
$$\Delta_2 : f_i - f_{i+1}$$

MEDIDAS DE DISPERSIÓN

3.1 Varianza

$$\sigma^2 = \sum_{i=1}^{N} \frac{(x_i - \overline{x})^2}{N} f_i$$

$$\sigma^2 = \sum_{i=1}^{N} \frac{x_i^2 f_i}{N} - \sum_{i=1}^{N} \left(\frac{x_i f_i}{N} \right)^2$$

3.2 Cuasivarianza

$$S^2 = \sum_{i=1}^{N} \frac{(x_i - \overline{x})^2}{N - 1}$$

3.3 Coeficiente de Variación

$$CV = \frac{\sigma_x}{x} \times 100\%$$

3.4 Intervarianza

$$\begin{aligned} \sigma_b^2 &= V(\overline{x}_b) \\ &= \sum_{n-1}^{l} \frac{(\overline{x}_h - \overline{x}) \times N_h}{N} \end{aligned}$$

3.5 Intravarianza

$$\begin{aligned} \sigma_w^2 &= M\left(\sigma_h^2\right) \\ &= \sum_{n-1}^{l} \frac{\sigma_n^2 \times N_h}{N} \end{aligned}$$

3.6 Asimetría

$$As = \frac{3(\overline{x} - Med)}{s}$$

FUNCIÓN GENERADORA DE MOMENTOS

4.1 Densidad

$$f(x) > 0$$

$$\int_a^b f(x)dx = 1$$

4.2 Esperanza

$$E(x) = \int_{-\infty}^{\infty} x f(x)dx$$

4.3 Varianza

$$\begin{aligned} \sigma^2 &= E\left[(x-\mu)^2\right] \\ &= \sum (x-\mu)^2 f(x) \end{aligned}$$

4.4 Momentos

$$\begin{aligned} \mu_{k,a} &= E\left[(x-a)^k\right] \\ &= \int \epsilon^{(x-a)^k} f(x)\, dx \end{aligned}$$

Distribuciones de Probabilidad

5.1 Distribuciones de Probabilidad Discretas

5.1.1 Distribución de Bernulli

La distribución de Bernulli es una distribución de probabilidad discreta en el que existen solo probabilidad de éxito (**p**) y de fracaso (**q**)

$$0 < p < 1,\ p \in \mathbb{R}\ \wedge\ q = 1 - p$$

Sea **X** una variable aleatoria, que mide el número de éxitos cuando se realiza un único experimento con dos resultados posibles, entonces X se distribuye como una Bernulli de parámetro **p**.

$X \sim Be(p)$ con función de probabilidad $f(x) = p^x(1-p)^{1-x}$

$$\begin{aligned} \mu := E(X) &= 0 \times q + 1 \times p \\ &= p \end{aligned} \qquad \begin{aligned} \sigma^2 := var(X) &= p - p^2 \\ &= p(1-p) \end{aligned}$$

5.1.2 Distribución Binomial

Sea **X** una variable aleatoria que mide el número de éxitos en **n** experimentos con dos resultados posibles **p** y **q**, entonces X se distribuye como una Binomial de parámetros n y p.

$$0 < p < 1,\ p \in \mathbb{R}\ \wedge\ q = 1 - p\ ;\ n \in \mathbb{N}$$

$X \sim B(n,p)$ con función de robabilidad $f\ (x) = \binom{n}{x} p^x (1-p)^{1-x}$

$$\binom{n}{x} = \frac{n!}{x!\ (n-x)!}$$

$$\mu := E(X) = np \qquad \sigma^2 := var(X) = np(1-p)$$

n	x	p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1	0		0.9900	0.9800	0.9700	0.9600	0.9500	0.9400	0.9300	0.9200	0.9100
	1		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0		0.9801	0.9604	0.9409	0.9216	0.9025	0.8836	0.8649	0.8464	0.8281
	1		0.9999	0.9996	0.9991	0.9984	0.9975	0.9964	0.9951	0.9936	0.9919
	2		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0		0.9703	0.9412	0.9127	0.8847	0.8574	0.8306	0.8044	0.7787	0.7536
	1		0.9997	0.9988	0.9974	0.9953	0.9928	0.9896	0.9860	0.9818	0.9772
	2		1.0000	1.0000	1.0000	0.9999	0.9999	0.9998	0.9997	0.9995	0.9993
	3		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0		0.9606	0.9224	0.8853	0.8493	0.8145	0.7807	0.7481	0.7164	0.6857
	1		0.9994	0.9977	0.9948	0.9909	0.9860	0.9801	0.9733	0.9656	0.9570
	2		1.0000	1.0000	0.9999	0.9998	0.9995	0.9992	0.9987	0.9981	0.9973
	3		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	4		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0		0.9510	0.9039	0.8587	0.8154	0.7738	0.7339	0.6957	0.6591	0.6240
	1		0.9990	0.9962	0.9915	0.9852	0.9774	0.9681	0.9575	0.9456	0.9326
	2		1.0000	0.9999	0.9997	0.9994	0.9988	0.9980	0.9969	0.9955	0.9937
	3		1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9998	0.9997
	4		1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 5.1: Función de distribución acumulativa binomial

5.1.3 Distribución de Poisson

Sea **k** el número de ocurrencias del evento o fenómeno, y sea λ un parámetro positivo que representa el número de veces que se espera que ocurra el fenómeno durante un intervalo de tiempo dado.

$$\begin{aligned} \lambda &\in \{0, \infty\} \\ k &\in \mathbb{N} \end{aligned}$$

$$\begin{aligned} \mathrm{e} &:= \lim{}_{n\to\infty}\left(1+\tfrac{1}{n}\right)^n \\ &= 2.7182818284\ldots \end{aligned}$$

X la variable aleatoria que mide la probabilidad de dicho suceso, entonces X tiene una distribución Poisson.
$X \sim Poi(\lambda)$ con funcion de probabilidad $f\left(x\right) = \frac{e^{-x}\lambda^x}{x!}$

$$\begin{aligned} \mu &:= E(X) \\ &= \lambda \end{aligned} \qquad \begin{aligned} \sigma^2 &:= var(X) \\ &= \lambda \end{aligned}$$

5.1.4 Distribución Hipergeométrica

Si **N** es el tamaño de la población, **n** el tamaño de la muestra extraída, m el número de elementos de la población original deseada, **X** es el número de elementos en la muestra que pertenecen dicha categoría, Se tiene una distribución Hipergeométrica.

$$X \sim H(N, m, n) \text{ con función de probabilidad } f\left(x\right) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

$$\begin{aligned} \mu &:= E(X) \\ &= \frac{nm}{N} \end{aligned} \qquad \begin{aligned} \sigma^2 &:= var(X) \\ &= n\frac{m}{N}\left(\frac{N-m}{N}\right)\left(\frac{N-m}{N-1}\right) \end{aligned}$$

5.2 Distribuciones de Probabilidad Continuas

5.2.1 Distribución Normal

Sea **X** es una variable aleatoria continua se dice que X sigue una ley normal con media μ y varianza σ^2

$X \sim N(\mu, \sigma^2)$ con función de probabilidad $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}$

ESTIMACION DE INTERVALO PARA MEDIA POBLACIONAL

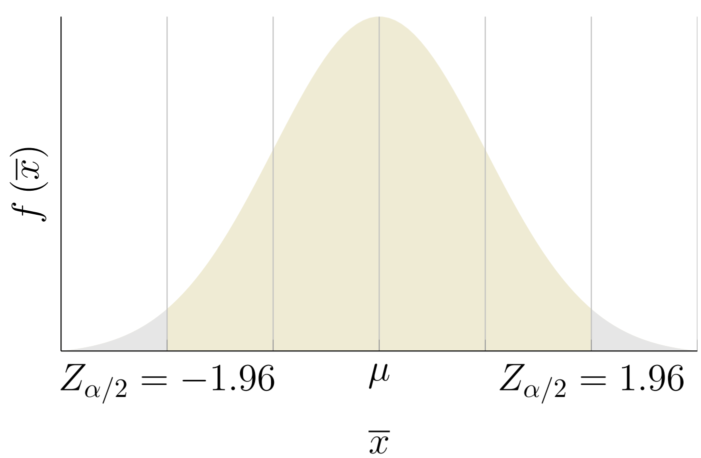
6.1 Varianza Conocida

$$\text{Población}\begin{cases} N \sim Normal\left(\mu\,;\,\sigma^2\right) \\ \mu : Desconocida \\ \sigma^2 : Conocido \end{cases} \qquad \text{Muestra}\begin{cases} n: \quad x_1, \quad x_2, \quad x_3, \quad \ldots \quad x_n \\ \overline{x} \\ s^2 \end{cases}$$

$$X \sim Z \sim N\left(\mu;\sigma^2\right)$$

$$Z = \frac{x-\mu}{\sigma\sqrt{n}}$$

$$P\left[\overline{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{n} \ < \ \mu \ > \ \overline{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{n}\right] = 1-\alpha$$



6.2 Varianza Desconocida

$$\text{Población}\begin{cases} N \sim Normal\left(\mu\,;\sigma^2\right) \\ \mu : Desconocida \\ \sigma^2 : Desconocida \end{cases} \qquad \text{Muestra}\begin{cases} n : \quad x_1, \quad x_2, \quad x_3, \quad \ldots \quad x_n \\ \overline{x} \\ s^2 \end{cases}$$

$$X \sim Z \sim t_{(v,\frac{\alpha}{2})}$$

$$Z = \frac{x - \mu}{\sigma\sqrt{n}}$$

$$P\left[\overline{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{n} \; < \; \mu \; > \; \overline{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{n}\right] = 1 - \alpha$$