UNAP
F I N E S I

IV Semestre

Modelos Discretos

Datos de Respuesta categóricos

Mtr. Alcides Ramos Calcina

Datos de Respuesta Categóricos

Mtr. Alcides Ramos Calcina

Datos de Respuesta Categóricos

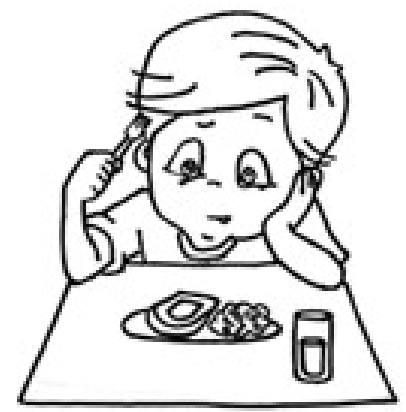
Una variable categórica tiene una escala de medición que consta de un conjunto de categorías. Por ejemplo,

Variable:
Ideoloía política

Categorías:

- Izquierda
- Centro
- Derecha

Variable:
Diagnósticos de anemía en niños



Categorías:

- Leve
- Moderada
- Grave

Datos de Respuesta Categóricos

Los datos categóricos no están de ninguna manera restringidos a las ciencias sociales y biomédicas. Ocurren con frecuencia en otras áreas, por ejemplo:

- Ciencias del comportamiento

  Variable: Tipo de enfermedad mental

  Categorías: Esquizofrenia, depresión, neurosis

- Epidemiología y salud pública

  Varibale: Método anticonceptivo en la última relación sexual

  Categorías: ninguno, condón, píldora, DIU, otro

- Genética

  Variable: Tipo de alelo heredado por una descendencia

  Categorías: Alelo recesivo, dominante.

- Botánica y zoología

  Variable: Presencia de organismo particular en un cuadrante muestreado

  Categorías: Si o no

- Educación

  Variable: Respuesta de un estudiante a una pregunta de examen

  Categorías: Correcta o incorrecta

Variable de Respuesta y Explicativa

En un análisis estadístico se distinguen entre.

- Variables de respuesta (o dependientes)
- Variables explicativas (o independientes).

En este curso nos centramos en métodos para variables de respuesta categóricas, como es el caso del modelo de regresión logístico y donde las variables explicativas pueden ser de cualquier tipo.

Función logística

$$p(x)=\frac{e^{f(x)}}{1+e^{f(x)}}=\frac{e^{\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_kx_k}}{1+e^{\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_kx_k}}=\frac{1}{1+e^{-\beta_0-\sum_{i=1}^{k}\beta_ix_i}}$$

Ejemplo.

$$p(x)=\frac{1}{1+e^{[1.436+1.827x_{11}+2.993x_{12}+3.222x_3+20.631x_{32}+3.364x_{33}+21.069x_{34}+0.00x_4-2.816x_6]}}$$

Escala de medición de la Variable Respuesta

Variable Estadística

Es una característica observable de la población que se va investigar y que puede tomar diferentes valores. Generalmente son representados matemáticamente por $X_1, X_2, \ldots, X_n$.

Dato

Es un valor particular de la variable.

En los individuos de la población puneña, de uno a otro es variable:

- El grupo sanguíneo: {A, B, AB, O}
- Su nivel de felicidad: {deprimido, …, muy feliz}
- El número de hijos: {0, 1, 2, 3, ….}
- La estatura: {1.62, 1.75, 1.89, 152, ….}
- Condición laboral: {Empleado, desempleado}
- Salario: {1800, 800, 1100, …..}

Clasificación de Variable Estadistica

CUALITATIVO

Se denominan así a las variables que se refieren a cualidades o aquellas que solo pueden describirse.

Nominal

Es aquella que surge cuando se definen niveles o categorías, que no tienen ningún orden, en las distintas categorías existentes.

Ejemplo:

- Clasificación de un grupo de estudiantes por sexo: (masculino y femenino)
- Tipo de ideología económica: {capitalismo, socialismo, economía mixta, …}
- Nacionalidad: {peruano, argentino, mexicano, …}

Ordinal

Variable que surge cuando el investigador ordena los datos de acuerdo con cierto criterio jerárquico y conservando un orden en particular.

Ejemplo:

- Rendimiento académico de los estudiantes de la UNA {excelente, bueno, regular, deficiente}
- Cadena de mando: {empleado, jefe, director, …, gerente}.
- Calificación crediticia (mala, regular, buena)

Clasificación de Variable Estadística

CUANTITATIVO

Se denominan así a las variables que proceden de procesos de medición o conteo.

Discreto

Se denomina así, por que éstas toman valores enteros o de un conjunto contable (numérico o finito).

Ejemplo:

- Número de empleados por área
- Número de estudiantes por semestre
- Número de proyectos en ejecución por municipio.
- Número de Pymes por región.

Continuo

Se denomina así a las variables que toman valores reales en un intervalo, de acuerdo a la naturaleza de la variable.

Ejemplo:

- Temperatura del medio ambiente: 8.5º, 17.3, 30.1, etc.
- Peso de una bolsa de arroz (kg): 1.25, 1.06, 1.11, …, etc.
- Ingreso mensual por trabajador: S/.1820.56, S/.1220.10, S/. 3000.0, etc.

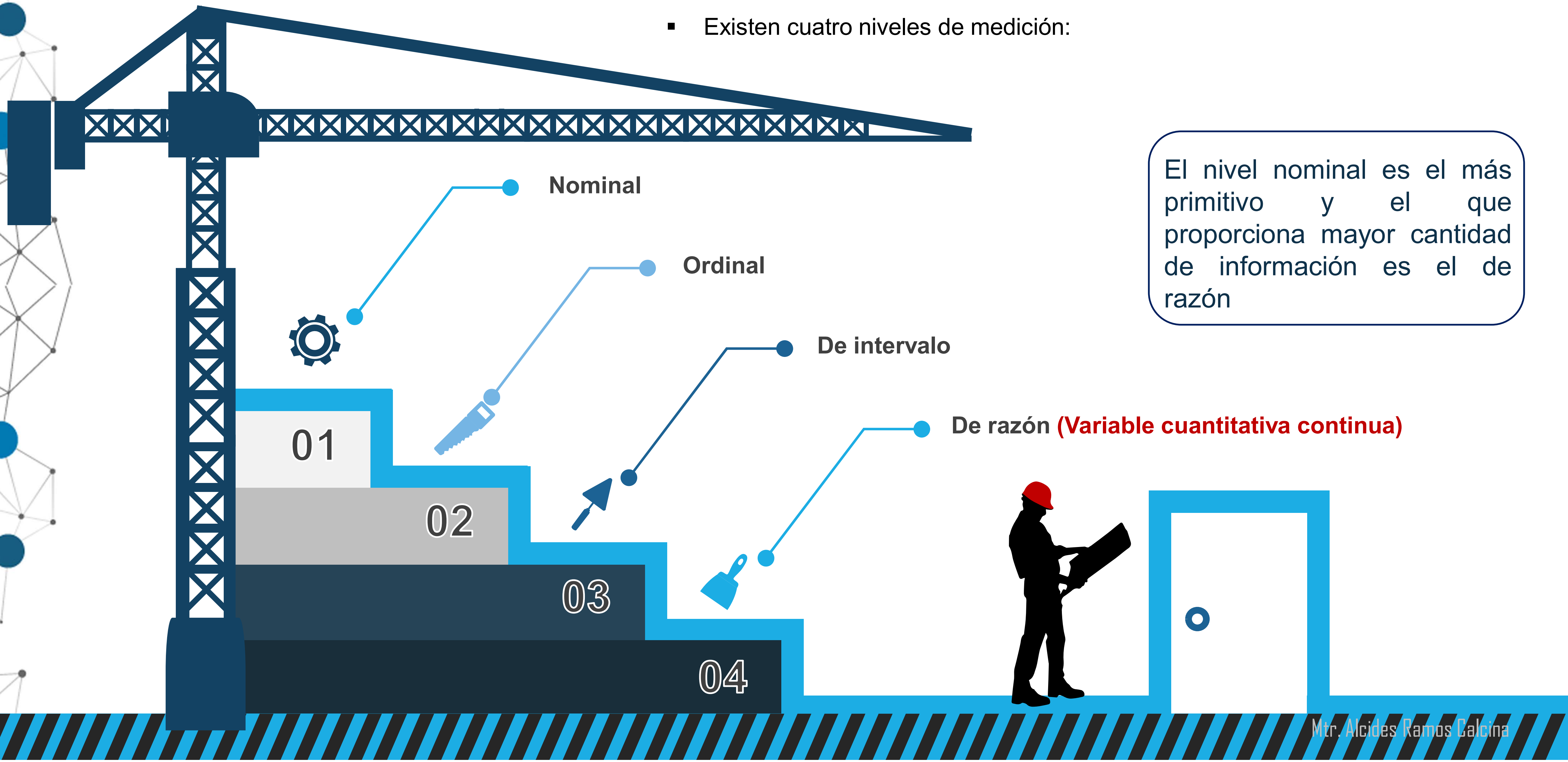
Escalas de **Medición**

Definición

- Una escala es un patrón convencional de medición, y básicamente consiste en un instrumento capaz de representar con gran fidelidad verbal, gráfica o simbólicamente el estado de una variable.

Escalas de Medición

- Existen cuatro niveles de medición:



El nivel nominal es el más primitivo y el que proporciona mayor cantidad de información es el de razón

Escala Nominal

- Las variables nominales tienen dos reglas:

Dicotómicas	Policotómicas
• Las categorías diferencian una forma de otra y son mutuamente excluyentes. • **Ejemplo**: El personal de una empresa: gorda o flaca, varón o mujer, alto o bajo, ingeniero o técnico, etc.	• Las categorías de la variable deben ser exhaustivas, es decir, incluir todas las posibles alternativas de variación de la variable. • **Ejemplo**: Tipos de sangre, incluye las siguientes categorías: O, A, B, AB. No existe otra categoría.

Escala Ordinal

- Las variables ordinales indican categorías que guardan un orden jerárquico.
- Requieren nivel de ordenación en su clasificación.

Ejemplo

Nivel de avance del proyecto	: baja, moderada, avanzada
Nivel educativo	: iletrado, primaría, secundaría, superior
Nivel de satisfacción	: bajo, medio, alto
Rendimiento académico	: superior, promedio, inferior

Escala de Intervalo

- Son categorías formadas por características numéricas con un orden jerárquico.
- Tienen todas las propiedades de las variables nominales y ordinales, con algo adicional: números que miden la distancia entre cada categoría.
- Los números en estas variables pueden ser continuos o discretos.

Ejemplo

Intervalo discreto:

Grupo de edad:

[ ] menores a 1 año
[ ] 1 a 5 años
[ ] 6 a 10 años
[ ] 11 a 15 años
[ ] 16 a más

Intervalo continuo:

Peso de materiales:

[20.0 – 25.5 〉
[25.5 – 30.0 〉
[30.0 – 35.5 〉
[35.5 – 40.0 〉
[40.0 – 45.5 〉

Escalas de Actitud

Opinión

Es una postura más estática, representa una posición mental consciente y manifiesta sobre algo o alguien. No implica disposición a la acción.

Actitud

Es un estado de disposición psicológica, adquirida y organizada a través de la propia experiencia que incita al individuo a reaccionar de una manera característica frente a determinados estímulos.

Escalas de Actitud

Pruebas o Escalas de actitud

Evalúan opiniones

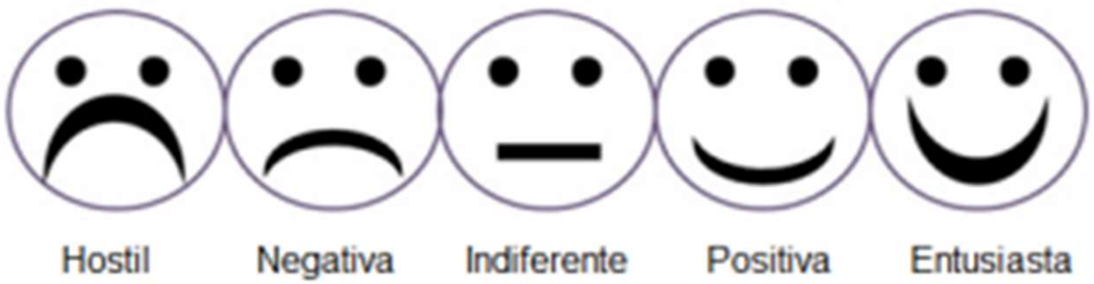
- Objeto
- Persona
- Suceso

Mtr. Alcides Ramos Calcina

Escalas de Actitud

Escala de Likert

- En este tipo de escalas se ofrece una afirmación al sujeto y se pide que la califique del ***0 al 4*** según su grado de acuerdo con la misma.
- Estas afirmaciones pueden reflejar actitudes positivas hacia algo o negativas.
- Las primeras se llaman favorables y las segundas desfavorables.
- Es muy importante que las afirmaciones sean claramente positivas o negativas, toda afirmación neutra debe ser eliminada
- Es una escala aditiva que corresponde a un nivel de medición ordinal.



Escala de Likert

Ejemplo:

- Señala tu grado de acuerdo o desacuerdo con la siguiente afirmación: "***La exposición fue un éxito***".

4 ( ) Totalmente de acuerdo
3 ( ) De acuerdo en general
2 ( ) Ni de acuerdo ni en desacuerdo
1 ( ) En desacuerdo en general
0 ( ) Totalmente en desacuerdo

Si los ítems son negativos:

0 ( ) Totalmente de acuerdo
1 ( ) De acuerdo en general
2 ( ) Ni de acuerdo ni en desacuerdo
3 ( ) En desacuerdo en general
4 ( ) Totalmente en desacuerdo

Escala de Gutman

- Soluciona el problema de la ambigüedad ya que es unidimensional.
- Es del tipo acumulativo.
- Se busca coherencia en las respuestas de los sujetos y esa coherencia es garantizada por el coeficiente de reproductividad.
- Se le presenta al sujeto una serie de cuestiones jerarquizadas de mayor a menor y se pide su veracidad en cada caso.


Alegre


Triste


Bueno


Malo

Tipo de escala de medición

- **Escalas descriptivas o verbales.** Los rangos o escalas se manifiestan por medio de expresiones verbales más o menos descriptivas.

Categoría: Utiliza el equipo de medición correctamente:
Rangos: ( ) Nunca ( ) Pocas veces ( ) A veces ( ) Casi siempre ( ) Siempre

- **Escalas descriptivas con información adicional.** Los rasgos o escalas se manifiestan por medio de expresiones verbales con una descripción más amplia, relacionado con el rasgo y la actividad efectuada.

Categoría: Utiliza el equipo de medición correctamente:
Rangos:

( ) Nunca, sólo sus compañeros realizan las mediciones.
( ) Pocas veces solicita a sus compañeros realizar las mediciones.
( ) A veces solicita a sus compañeros realizar las mediciones.
( ) Casi siempre realiza las mediciones.
( ) Siempre es el primero en realizar las mediciones.

Tipo de escala de medición

- **Escalas numéricas**. Los rangos o escalas se manifiestan por medio de números asignados al grado de magnitud (fuerza) de la conducta. Pueden ser crecientes (mínimo a máximo) o decreciente (máximo a mínimo).

Categoría: Utiliza el equipo de medición correctamente:
Rangos: ( )1 ( )2 ( )3 ( )4 ( )5

- **Escalas gráficas**. Los rangos se manifiestan por medio de líneas o barras que sitúan las frases descriptivas o valores numéricos otorgando una mayor precisión en la selección del rango.

Categoría: Utiliza el equipo de medición correctamente:
Rangos: 1 2 3 4 5 6 7 8 9

DISTRIBUCIONES PARA DATOS CATEÓRICOS

Distribución Binomial

- Simbolizamos con ***n*** a la cantidad de veces que se repite el experimento.
- La variable la simbolizamos con **X** y se define como X: cantidad de éxitos ocurridos en las ***n*** repeticiones del experimento.
- Simbolizando con Rx al recorrido (o campo de variación) de la variable, tenemos que:

$$Rx = \{0 , 1 , 2 , \ldots , n\}$$

- Supongamos una serie de ***n*** repeticiones de pruebas Bernoulli, en la que los *éxitos y fracasos* resultaron:

$$\underbrace{EE.....EEE}_{k \text{ veces}}\ \underbrace{FF.....FFFFF}_{(n-k) \text{ veces}}$$

Distribución Binomial

- La probabilidad de que el resultado de las ***n*** repeticiones sea esta secuencia es:

$$P(\underbrace{EE.....EEE}_{k \text{ veces}}\,\underbrace{FF.....FFFFF}_{(n-k) \text{ veces}}) = \underbrace{P(E)P(E).....P(E)}_{k \text{ veces}}\,\underbrace{P(F)P(F).....P(F)}_{(n-k) \text{ veces}}$$

donde P(E) = p y P(F) = q

$$P(\underbrace{EE.....EEE}_{k \text{ veces}}\,\underbrace{FF.....FFFFF}_{(n-k) \text{ veces}}) = \underbrace{pp.....ppp}_{k \text{ veces}}\,\underbrace{qq.....qqqqq}_{(n-k) \text{ veces}} = p^k q^{n-k}$$

- Esta secuencia que hemos considerado es una de aquellas en las que aparece k veces éxito. Pero no estamos interesados en una secuencia particular de *éxitos* y *fracasos*, sino en la cantidad de *éxitos*, no importando el orden.

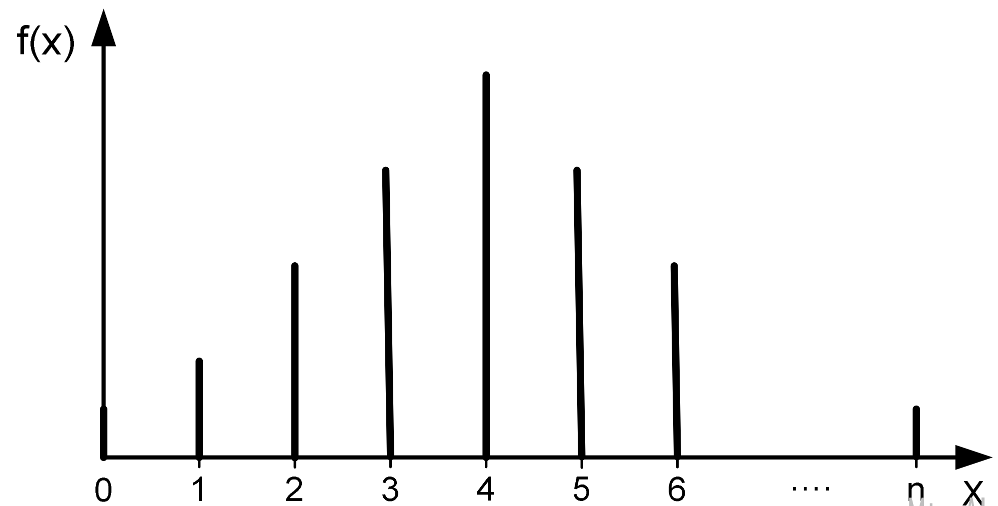
Distribución Binomial

DEFINICIÓN

- La variable aleatoria X = "número de éxitos en n experimentos" se dice que sigue una distribución binomial con parámetros *n* y *p*, X~B(*n*,*p*), si su función de probabilidad esta dado por:

$$P(x) = P[X = x] = \begin{cases} C_x^n p^x q^{n-x} & x = 0,1,2,\ldots,n \\ 0 & \text{otro caso} \end{cases}$$

- Gráfica de la función de probabilidad.



Distribución Binomial

Medidas Estadísticas

- Si una variable aleatoria X~B(n,p) con parámetros n y p, entonces:

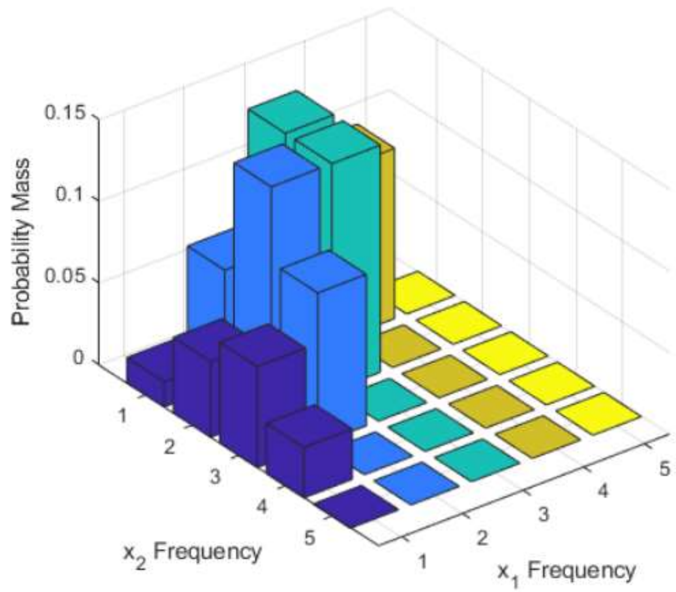
  - Media : $\mu = E[x] = np$
  - Varianza : $\sigma^2 = Var[X] = npq$

Distribución Multinomial

La distribución multinomial es una generalización de la distribución binomial, la cual se presentará cuando el experimento aleatorio en cuestión no sólo dé lugar a dos posibles resultados, éxito y fracaso, como ocurría en la binomial, sino que dé lugar a tres o más posibles resultados.

CARACTERÍSTICAS

- Se esperan más de dos tipos de resultados.
- Las probabilidades asociadas a cada uno de los resultados son constantes.
- Las repeticiones del experiment son independientes
- El número de repeticiones, *n,* es constante.



Distribución Multinomial

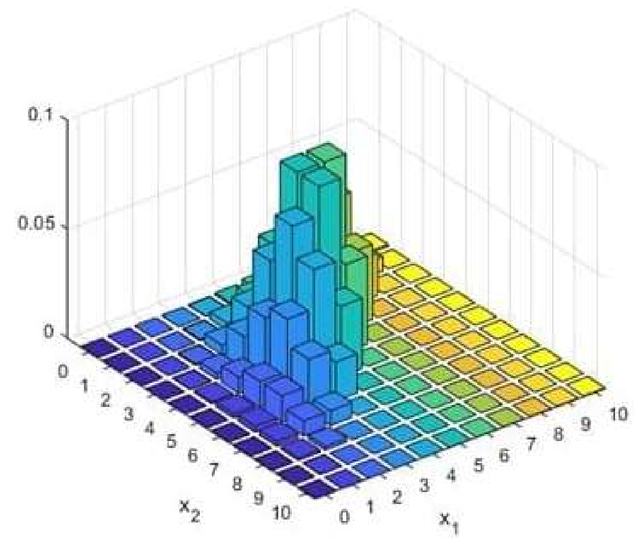
DEFINICIÓN

Decimos que la v.a. r-dimensional $(X_1, X_2, \ldots, X_r)$ = “número de veces que se presentan cada uno de los sucesos $A_1, A_2, \ldots, A_r$ cuando se realizan las n-repeticiones independientes del experimento” sigue una distribución multinomial de parámetros n, $p_1, p_2, \ldots, p_r$, $(X_1, X_2, \ldots, X_r) \sim M(n, p_1, p_2, \ldots, p_r)$ y su función de probabilidad es:

$$P(X_1 = k_1, X_2 = k_2, \ldots, X_r = k_r) = \frac{n!}{k_1! \cdots k_r!} p_1^{k_1} \cdot p_2^{k_2} \cdots p_r^{k_r}$$

Con $k_i = 0, 1, 2, \ldots, n$ y $\sum_{i=1}^{r} k_i = n$

Distribución multinomial para X1 y X2.



Distribución Multinomial

Medidas Estadísticas

Si X = (X1, X2, …, Xr) es una variable aleatoria con distribución multinomial con parámetros n y p = (p1, p2, …, pr), entonces:

Media : $\mu = \mathrm{E}\left[\mathrm{X_i}\right] = \mathrm{np_i}$

Varianza: $\sigma^2 = \mathrm{Var}\left[\mathrm{X_i}\right] = \mathrm{np_i}\left(1 - \mathrm{p_i}\right)$ con i = 1, 2, 3, …., r

DISTRIBUCIONES PARA DATOS CATEÓRICOS

Distribución de Poisson

- La variable Poisson se define como X: número de acontecimientos ocurridos en un determinado intervalo, (de tiempo, de superficie, de volumen, etc.), y puede asumir cualquier número no negativo, es decir:

$$x = 0, 1, 2, 3, ..., \infty$$

- En base a las tres hipótesis consideradas se demuestra (mediante el Análisis Matemático Diferencial) que la probabilidad de que ocurran k acontecimientos en un intervalo está dado por la siguiente expresión:

$$P(x = k) = \frac{e^{-\lambda}.\lambda^{k}}{k!}$$

donde λ representa la tasa de ocurrencia de ese fenómeno en dicho intervalo. Obviamente, el valor de λ depende de la magnitud del intervalo, siendo proporcional al mismo ya que la tasa de ocurrencia es constante.

Distribución de Poisson

DEFINICIÓN

- La variable aleatoria X: "número de veces que ocurre un evento por unidad de tiempo" decimos que sigue una distribución de Poisson de parámetro λ (λ>0), denotado por X~P(λ), si su función de probabilidad esta dado por:

$$P(x) = P[X = x] \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & \text{si } x = 0, 1, 2, \ldots. \\ 0 & \text{otro caso} \end{cases}$$

siendo e = 2.71828…

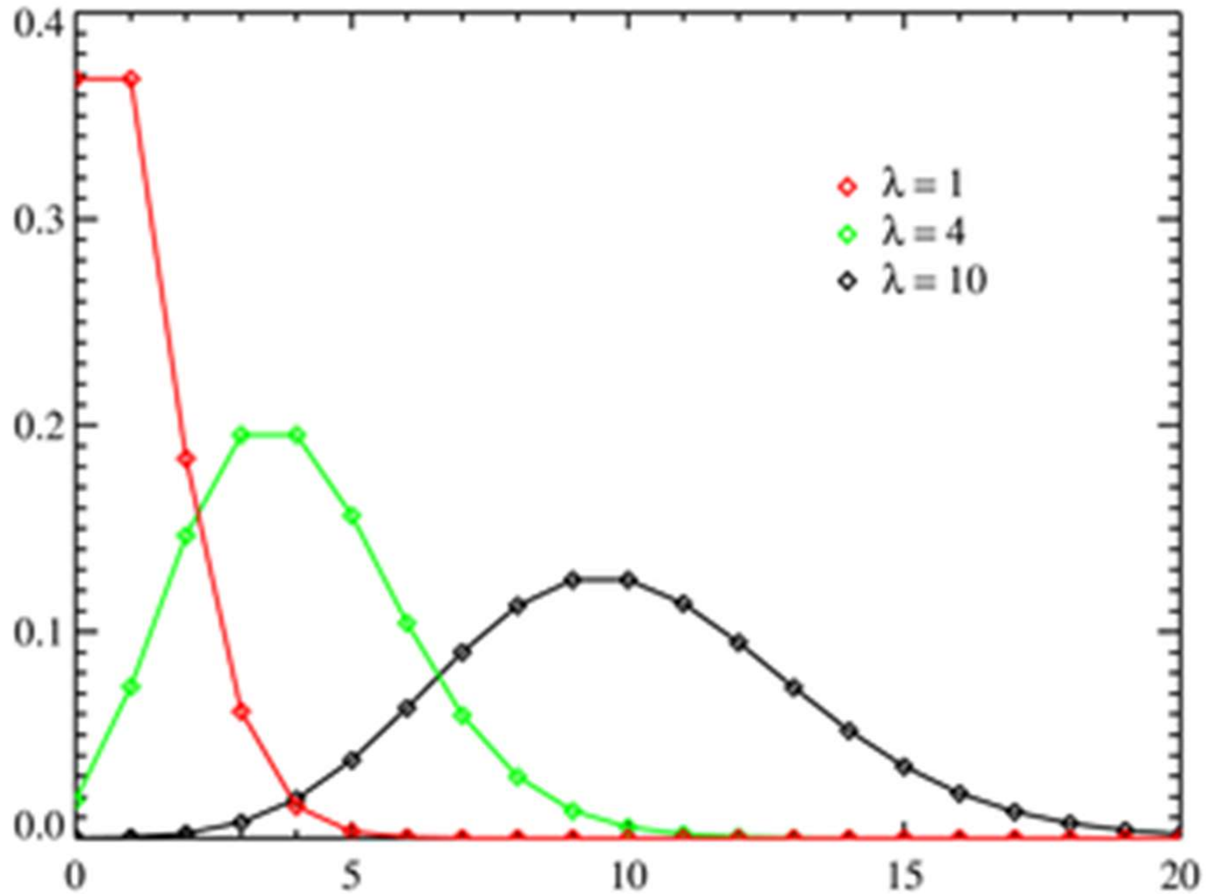
Verifiquemos, en primer lugar, que se trata de una verdadera distribución de probabilidad. Para ello calculamos

$$\sum_{x=0}^{\infty} P[X = x] = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^{2}}{2!} + \frac{\lambda^{3}}{3!}\cdots\right) = e^{-\lambda}e^{\lambda} = 1$$

Distribución de Poisson

Gráfica de la Función de Probabilidad

- Gráficamente la función de probabilidad de Poisson para diferentes valores de λ es.



Distribución de Poisson

Medidas Estadísticas

- Si X es una variable aleatoria con distribución de Poisson con parámetro $\lambda > 0$, entonces:

  - Media $\mu = \mathrm{E}[\mathrm{X}] = \lambda$

  - Varianza $\sigma^2 = \mathrm{Var}[\mathrm{X}] = \lambda$

Función Generatriz de Momentos

Su expresión será:

$$\varphi(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{e^{tx}\lambda^{x}}{x!} =$$

$$= e^{-\lambda}\left[1 + \frac{\lambda e^{t}}{1!} + \frac{(\lambda e^{t})^2}{2!} + \frac{(\lambda e^{t})^3}{3!} + \cdots\right] =$$

FINESI

Modelos Discretos

IV Semestre

GRACIAS

https://aulavirtual2.unap.edu.pe/