Datos bivariados

Tabla de registro de datos bivariados

- Es una tabla de datos compuesta por 3 columnas, la primera siendo los datos en análisis y las otras dos siendo las variables

	X	Y
Sujeto 1	x_1	y_1
$\vdots$	$\vdots$	$\vdots$
Sujeto i	x_i	y_i
$\vdots$	$\vdots$	$\vdots$
Sujeto n	x_n	y_n

X e Y. *

Tabla de frecuencia bidimensional

- Es una tabla que resume los datos de dos variables simultáneamente, mostrando la frecuencia.

Variable Y /Variable X	Categoría 1	Categoría 2	…	Categoría n	Total fila (Frecuencia absoluta X)
Categoría 1	*n11*	*n12*	…	*n1p*	$\sum_{i=1}^{p} n_{1i}$
Categoría 2	*n21*	*n22*	…	…	$\sum_{i=1}^{p} n_{2i}$
…	…	…	…	…	…
Categoría k	*nk1*	*nk2*	…	*nkp*	$\sum_{i=1}^{p} n_{ki}$
Total columna (Frecuencia absoluta de Y)	$\sum_{i=1}^{p} n_{i1}$	$\sum_{i=1}^{p} n_{i2}$	…	$\sum_{i=1}^{p} n_{ip}$	*n*

Frecuencia absoluta conjunta

- Pertenece a la clase conjunta de X_iY_j .
- Se denota por n_{ij}

- La suma de todas las frecuencias absolutas conjuntas debería ser la cantidad de datos n, esto es:

$$\sum_{i=1}^{p}\sum_{j=1}^{k} n_{ij} = n$$

Frecuencia relativa conjunta

- Pertenece a la clase conjunta de $X_i Y_j$ respecto a la cantidad de datos n.
- Se denota por f_{ij}
- la formula es:

$$f_{ij} = \frac{n_{ij}}{n}$$

- También la suma de las frecuencias relativas conjuntas debería ser 1.

$$\sum_{i=1}^{p}\sum_{j=1}^{k} f_{ij} = 1$$

- Si es expresada en porcentajes el resultado tiene que ser 100%.

Observaciones

1) Si ambos datos son de tipo cualitativo la tabla bidimensional se llama “tabla de contingencia”.
2) La tabla bidimensional también puede ser representada por frecuencia relativa conjunta o frecuencia relativa conjunta porcentual
3) A partir de la distribuciones bidimensional se pueden obtener distribuciones unidimensionales como la tabla marginal y/o condicionada.

Distribuciones marginales

Muestra como se distribuye una de las variables, ya sea X o Y, sin

Distribución Marginal de X:

X	n_i	f_i
x_1	n_1	f_1
x_2	n_2	f_2
$\vdots$	$\vdots$	$\vdots$
x_p	n_p	f_p
Total	n	1

tomar en cuenta la otra.

Distribución Marginal de Y:

Y	n_i	f_i
y_1	n_1	f_1
y_2	n_2	f_2
$\vdots$	$\vdots$	$\vdots$
y_k	n_k	f_k
Total	n	1

Distribuciones condicionadas Muestra como se distribuye una de las vari-

Distribución Condicional de X dado $Y = y_j$:

X	n_{ij}	$f_{i/j}$
x_1	n_{1j}	$f_{1/j} = n_{1j}/n_{*j}$
x_2	n_{2j}	$f_{2/j} = n_{2j}/n_{*j}$
$\vdots$	$\vdots$	$\vdots$
x_p	n_{pj}	$f_{p/j} = n_{pj}/n_{*j}$
Total	n_{*j}	1

Distribución Con dado X

Y	n_{ij}	
y_1	n_{i1}	$f_{i/}$
y_2	n_{i2}	$f_{i/}$
$\vdots$	$\vdots$	
y_k	n_{ik}	$f_{i/}$
Total	n_{i*}	

ables X o Y en base a una condición dada.
las frecuencia se obtiene en base al total de datos de una variable.

Covarianza

La covarianza se define como:

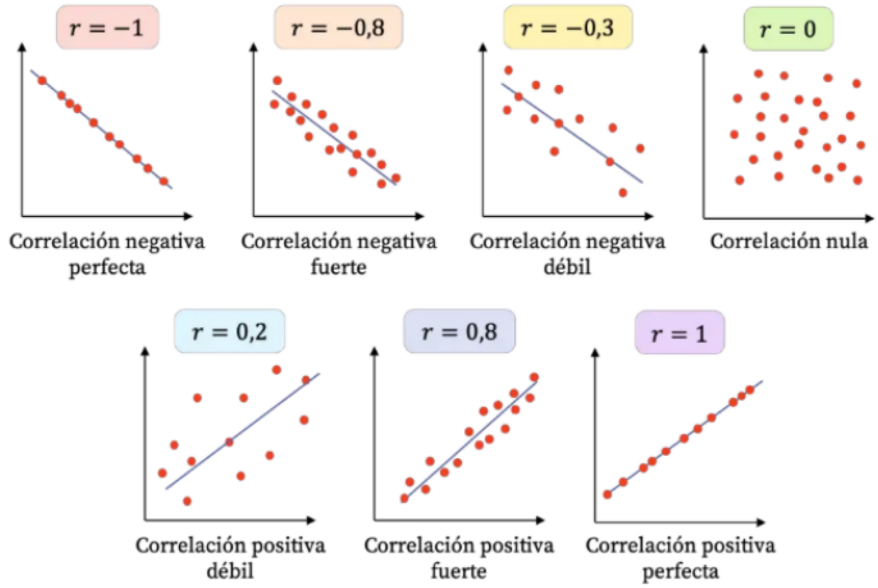
$$S_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{n-1}$$

- si $S_{xy} > 0$ entonces ambas variables aumenta simultáneamente.
- si $S_{xy} < 0$ entonces una variable aumenta mientras la otra disminuye.
- si $S_{xy} = 0$ entonces no hay una relación lineal entre ambas variables o la tendencia positiva y negativa se contrarrestan.

Correlación se define como:

$$r = \frac{S_{XY}}{S_X S_Y}$$

El coeficiente de correlación de Pearson mide la asociación lineal entre dos variables, mide que tan bien se asocian en relación a una fun-



ción lineal las variables.
Modelo de regresión lineal simple Busca crear una función lineal que se ajuste lo mas posible a lo datos, buscando que los datos estén relativamente cerca de la función creada. Se define como:

$$Y = \beta_1 X + \beta_0$$

Para encontrar β_1:

$$\beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} x_i^2 - n\bar{X}^2}$$

para encontrar β_0:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

Probabilidades

Conteo de probabilidades: ### Caso 1 dos operaciones que no se pueden hacer juntas: En caso de que una operaciones se pueda hacer n_1 formas y también de n_2 formas, pero no se pueden hacer juntas entonces la cantidad de operaciones que se pueden hacer serán $n_1 + n_2$. Si se realizan k operaciones distintas que no pueden llevarse a cabo de manera simultanea, solo se pueden llevar de n_k formas, entonces la operación será: $n_1 + n_2 + ... + n_k$

Caso 2 dos operaciones que si por cada una se puede llevar acabo de otra forma:

Si una operación se puede realizar de n_1 formas y por cada una des estas se puede realizar también de n_2 forma, entonces la operación se podrá realizar de $n_1 * n_2$ formas.

Similar a lo anterior teniendo en cuenta de k casos entonces: $n_1 * n_2 * ... * n_k$

Caso 3 permutación lineal, el caso donde el orden importa:

En este caso estamos hablando de cuando tratamos de ordenar algo y que un cambio en este orden significa una disposición distinta ABC , no será lo mismo que BAC, la permutación de n objetos distinto será n!:

$$n! = n * (n-1) * (n-2)...1$$

Permutaciones circulares: El caso de las permutaciones circulares puede resultar un poco confuso, si por ejemplo tenemos a unas personas alrededor de una mesa redonda, se considerara que es una posición distinta cuando sus vecinos cambien, no cuando roten de silla. su formula es $(n-1)!$

Permutaciones con elementos repetidos: Es un caso diferente, en este tomamos en cuenta cuando existen n objetos del tipo n_1 o de n_2 , la ecuación es:

$$\frac{n!}{n_1! * n_2!...n_k!}$$

Observaciones ==Hablamos de un experimento aleatorio cuando el resultado es incierto y se denota por ϵ y se denomina experimento determinístico cuando se sabe el resultado de antemano ==

==Una acción no es un experimento estadístico, lanzar una moneda es una acción, pero a nosotros nos interesa el resultado==

Espacio muestral: Es el conjunto de todos los posibles resultados asociados a un experimento. Su símbolo es Ω, si el espacio muestral tiene numero finito de elementos o infinitos numerables entonces se dice que es discreto, si el intervalo son todos los puntos de algún intervalo real, entonces es continuo.

Se denomina Evento o Suceso a cualquier subconjunto del espacio muestral y se denota por las letras mayúsculas.

Son sucesos:

1. Obtener un número par al hacer girar una rulet del 1 al 12.
$$A = \{2, 4, 6, 8, 10, 12\}$$
2. Obtener al menos una cara al lanzar dos monec
$$B = \{cs, sc, cc\}$$

En particular el propio espacio muestral es un evento y se llama suceso seguro, mientras que el conjunto vacío ϕ, es un evento y se llama suceso imposible.

Como los eventos son subconjuntos es posible aplicar la teoría de los conjuntos para obtener nuevos eventos, Si A y B son eventos, entonces también lo son A ∪ B, A ∩ B, A^c

Diremos que dos eventos A y B son mutuamente excluyentes o disjuntos si no pueden ocurrir juntos, es decir A ∩ B = ϕ.

Probabilidad: Si Ω es un espacio muestral con n elementos, entonces la probabilidad de que un evento A ocurra será:

$$P(A) = \frac{m}{n}$$

donde m es el numero de elementos de A. Se conoce como regla de Laplace y solo funciona en un espacio finito equiprobable.

Preposición:

1) $P(\Omega) = 1$
2) $P(A) \geq 0$
3) $P(A \cup B) = P(A) + P(B)$ Si $A \cap B = \emptyset$ ##### Teorema:
4) $P(\emptyset) = 0$
5) $P(A^c) = 1 - P(A)$
6) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Probabilidad Condicional: Cuando se calcula la probabilidad de un evento A en particular y se tiene información de la ocurrencia de otro evento B, se le conoce como probabilidad condicional, se denota por $P(A|B)$ se lee "Probabilidad de A dado B" y se define como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

con $P(B) \neq 0$

Example 18

Se realizó una encuesta sobre hábitos de lectura que se resume por medio de la tabla.

	Le gusta leer	No le gusta leer	Total
Hombre	10	20	30
Mujer	90	30	120
Total	100	50	150

Determine la probabilidad que si una persona elegida al azar le guste leer dado que es mujer.

Regla multiplicativa: **Eventos independientes:**

$$P(A \cap B) = P(A) * P(B)$$

Evento Dependiente:

$$P(A \cap B) = P(A) * P(B|A)$$

$$P(A \cap B) = P(B) * P(A|B)$$

Example 19

Dado ϵ : Lanzar una moneda tres veces. Determine la probabilidad de obtener cara en el primero, cara en el segundo y sello en el tercero.

Probabilidad total:

Teorema(probabilidad total) Sea $A_1, A_2, ..., A_n$ una partición de Ω 1) $\cup_{i=1}^{n} A_i = \Omega$ 2) $A_i \cap A_j = \emptyset$ Sea A otro evento de Ω, entonces

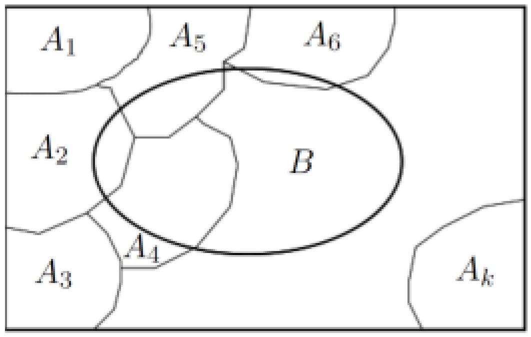
$$P(A) = \sum_{i=1}^{n} P(A|A_i) * P(A_i)$$

Example 21

Hay tres urnas, la primera tiene 2 bolas blancas y 2 negras, la segunda tiene 2 blancas y una negra, y la tercera tiene una bola blanca y 3 negras. Se extrae una bola al azar ¿Cuál es la probabilidad de que esa bola sea blanca?

Teorema de Bayes: Sea $A_1, A_2, ..., A_n$ una partición de Ω y B otro evento de Ω, entonces

$$P(A_i|B) = \frac{P(B|A_i) * P(A_i)}{\sum_{i=1}^{n} P(B|A_i) * P(A_i)}$$



Certamen 3

Variable aleatoria

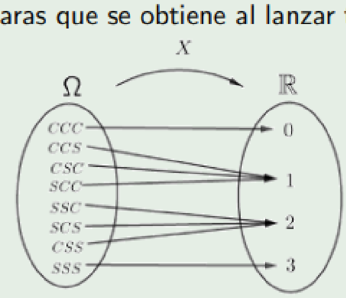
sea una variable aleatoria X en una función que asigna un número real a cada resultado en el espacio muestra Ω de un experimento aleatorio:

$$X : \Omega \rightarrow \mathbb{R}$$

$$w_i \rightarrow X(w_i) = x_i$$

Example 1

Sea X el número de caras que se obtiene al lanzar tres monedas.



Una vez definida la variable aleatoria tenemos que definir su rango, denotado por R_X , que son todos los valores posibles de X

Dependiendo cual es el recorrido numérica de X podemos agruparla en : **V.A. Discreta:** si el recorrido de X es un conjunto finito o infinito numerable. **V.A. Continua:** si el recorrido de X es un conjunto no

Example 2

Ejemplos **V.A. Discreta**:

- X: Suma de los números que aparecen en las caras cuando se lanzan dos dados (finito numerable).
- X: Número de lanzamientos de una moneda hasta que salga una cara (infinito numerable).

Ejemplos **V.A. Continua**:

- X: Tiempo (meses) de vida de un Smartphone.
- X: Largo (metros) de un cable del tendido eléctrico.

contable.

Función de probabilidad

Recibe el nombre de función de probabilidad o distribución de probabilidad la función $f_X(x) = P(X = x)$ que va desde el conjunto de valores [0,1], se tiene que cumplir las siguientes condiciones: 1. $0 \leq P(X = x) \leq 1$ 2. $\sum_{x \in R_X} P(X = x) = 1$

Función de densidad

En el caso de una variable aleatoria continua, no tiene sentido estudiar la probabilidad de valores puntuales porque da 0:

$$P(X = a) = \frac{1}{\infty} = 0$$

Por eso se estudia la función de densidad de probabilidad, denotada por f_X, que describe la densidad de probabilidad en torno a cualquier punto x, permite calcular la probabilidad en intervalos $P(a \leq x \leq b)$

En el caso de que X sea una v.a. continua, la función densidad de X, $f_X(x)$, es una función tal que: 1. $f_X(x) \geq 0 \forall x \in R_X$ 2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$ (Área bajo la curva)

Entonces la probabilidad de que ocurra un evento A lo calculamos así:

$$P(X \in A) = \int_A f(x)dx$$

Función de distribución acumulada

en caso discreto se define como:

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i)$$

Con $i \in R_X$ y cumple la siguiente propiedad

$$P(a \leq X \leq b) = P(X \leq b) - P(x \leq a)$$

Esperanza

describe el centro de la distribución de probabilidad de una variable aleatoria y es un valor promedio esperado a largo plazo si se repitiera el experimento aleatorio muchas veces.

Sea X un v.a. con función de probabilidad o función de densidad f(x). Sea g(X) una función de la v.a. X. El valor esperado de g(X), simbolizado por E[g(x)] es:

$$E\left[g(x)\right] = \begin{cases} \displaystyle\sum_{x \in R_x} g(x) \cdot f(x) & \text{Si } X \text{ es una v.a. discreta} \\ \displaystyle\int_{-\infty}^{\infty} g(x) \cdot f(x) dx & \text{Si } X \text{ es una v.a. continua} \end{cases}$$

Esperanza y Varianza

1. Si g(X) = X entonces si está calculando la Esperanza de la v.a. X.

$$E\left[X\right] = \begin{cases} \displaystyle\sum_{x \in R_x} x \cdot f(x) & \text{Si } X \text{ es una v.a. discreta} \\ \displaystyle\int_{-\infty}^{\infty} x \cdot f(x) dx & \text{Si } X \text{ es una v.a. continua} \end{cases}$$

2. Si $g(X) = (X - \mu)^2$ entonces E[g(x)] se llama varianza de la v.a. y se simboliza como σ^2 .

$$E\left[X\right] = \begin{cases} \displaystyle\sum_{x \in R_x} (X-\mu)^2 \cdot f(x) & \text{Si } X \text{ es una v.a. discreta} \\ \displaystyle\int_{-\infty}^{\infty} (X-\mu)^2 \cdot f(x) dx & \text{Si } X \text{ es una v.a. continua} \end{cases}$$

Aplicando propiedades de la esperanza se obtiene el siguiente resultado

$$Var(X) = \sigma^2 = E(X^2) - [E(X)]^2$$

así, $$\begin{cases} \displaystyle\sum_{x\in R_x} x^2 \cdot f(x) - \left[\sum_{x\in R_x} x \cdot f(x)\right]^2 & \text{Si } X \text{ es una v.a. discreta} \\ \displaystyle\int_{-\infty}^{\infty} x^2 \cdot f(x)dx - \left[\int_{-\infty}^{\infty} x \cdot f(x)dx\right]^2 & \text{Si } X \text{ es una v.a. continua} \end{cases}$$

Modelos de probabilidad

Distribución de Bernoulli

Es el experimento mas básico, solo tiene dos resultados, denominado éxito y fracaso es decir Ω = {éxito, fracaso}.

Cualquier experimento de bernoulli se puede denotar A como éxito y A^c como fracaso.

Depende de un solo parámetro p, probabilidad de éxito, y entonces 1-p es la probabilidad de fracaso(1-p=q), donde 0<P<1.

sea Ω el espacio muestral de un experimento, sea $A \subseteq \Omega$ cualquier evento con P(A) = p, 0<p<1 y sea X la v.a. definida por

$$X(x) = \begin{cases} 1 & x \in A \\ 0 & x \in A^c \end{cases}$$

entonces se llama v.a Bernoulli con parámetro p, y se denota

X~b(p) o bien X~ Ber(p)

Si el experimento de Bernoulli se repite n veces, podemos definir una variable aleatoria X como la cantidad de éxitos de un total de n, esto es un experimento binomial.

Sea X el número total de éxitos en un experimento binomial con n ensayos y parámetro p. Entonces se llama v.a. binomial con parámetro n y p y se denota por $X \sim B(n,p)$ o bien $X \sim Bin(n,p)$

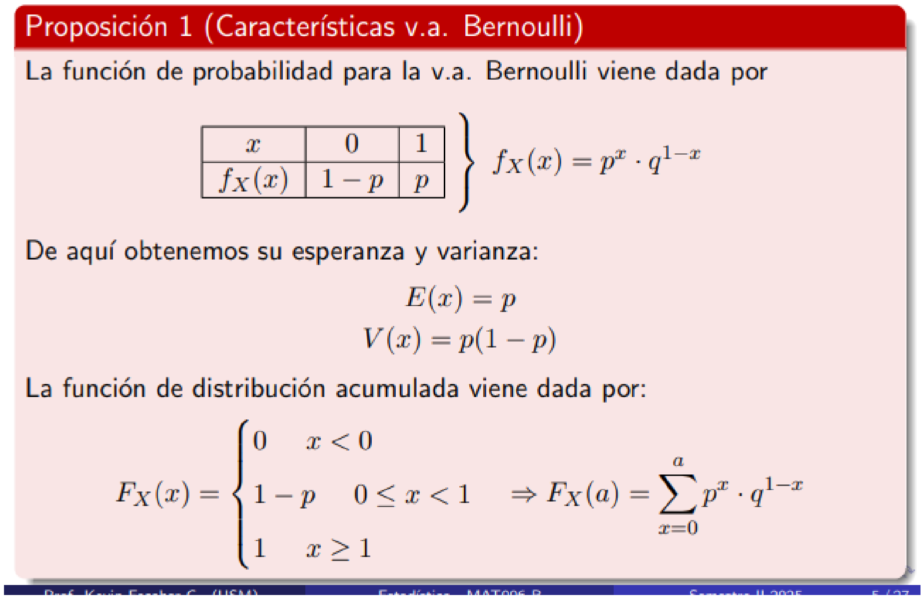
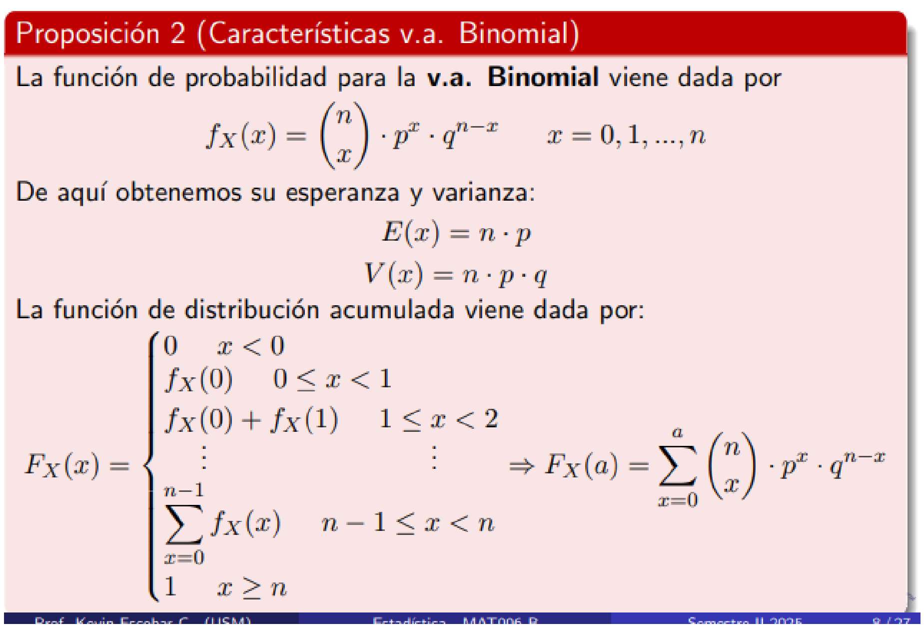


Figure 1: Pasted image 20251116165652.png



Distribución hipergeométrica

Proposición 2 (Características v.a. Hipergeométrica)

La función de probabilidad para la **v.a. Hipergeométrica** viene dada por

$$f_X(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \qquad x = 0, 1, ..., n$$

Donde

N:Tamaño total de la población.

k: Número total de éxitos en la población.

n: Tamaño de la muestra extraida.

x: Número de éxitos deseados en la muestra.

De aquí obtenemos su esperanza y varianza:

$$E(x) = n \cdot p \quad \text{donde} \quad p = \frac{k}{N}$$

$$V(x) = n \cdot p \cdot (1-p)\frac{(N-n)}{(N-1)}$$

Distribución Poisson

Los experimentos que resultan en valores numéricos de una variable aleatoria X y que representan el número de resultados durante un intervalo de tiempo dado o en una región específica frecuentemente se llaman experimentos Poisson.

Un experimento Poisson puede generar observaciones para una cierta v.a que representen el número de llamadas telefónicas por hora que se recibe en una oficina, el número de días en que una determinada escuela se cierra en invierno debido a la nieve, o al número de juegos pospuestos debido a la lluvia durante una temporada de fútbol.

La variable aleatoria Poisson es de la forma:

$$f_X(x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

De aquí tenemos su esperanza y varianza.

$$E(x) = \mu = \lambda$$

$$V(x) = \sigma^2 = \lambda$$

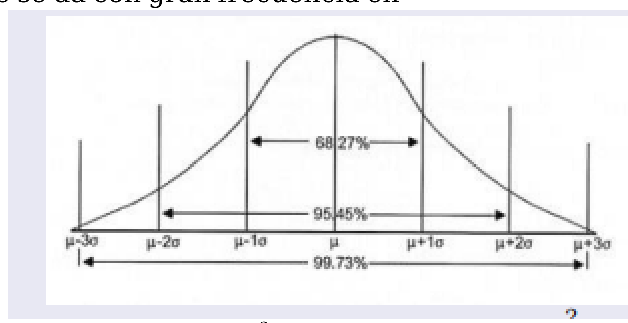
donde: λ : Tasa de ocurrencia promedio(por unidad de intervalo): numero de eventos ocurrido por unidad de tiempo o espacio(cliente por minutos)

t : Longitud del intervalo(tiempo o espacio): Define la duración o magnitud del intervalo (1 hora, 10 metros, etc).

λt : El parámetro de la distribución(μ o media): Frecuentemente a la combinación λt se le denota por una sola letra, ya sea μ o simplemente Λ. Este producto representa el número promedio total de eventos esperados en el intervalo total de la longitud t. se usa como media y varianza de la distribución.

Distribución normal

Una distribución de probabilidades que se da con gran frecuencia en



la naturaleza es la distribución normal.
Gauss plante que la función de densidad $f_X(x) = ke^{-\frac{x^2}{2}}$ cuyo núcleo esta dado por la expresión $e^{-\frac{x^2}{2}}$ (le da forma de campana) y la constante k la deduce de tal forma que el área bajo dicha curva sume 1, para que cumpla la función de densidad:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Sea X una v.a con función de densidad

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$$

Entonces, diremos que X tiene una distribución normal(Gaussiana) estándar y se denota

$$X \sim N(0,1)$$

Donde $\mu = 0$(media) y $\sigma^2 = 1$(varianza).

Sea X una variable aleatoria continua con función de probabilidad

$$f_X(s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, x \in \mathbb{R}, \mu \in \mathbb{R} y \sigma > 0$$

Entonces diremos que X tiene una distribución normal con parámetros μ(media) y σ^2 (varianza). Esto se denota

$$X \sim N(\mu, \sigma^2)$$

podemos notar que la distribución normal estandarizada no es mas que un caso particular de la distribución normal cuando $\mu = 0$ y $\sigma^2 = 1$

Si consideramos que $P(a < x < b) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$ esto no es sencillo por eso se usa esto: **Teorema(Estandarización)** sea X una variable aleatoria continua, luego

$$X \sim N(\mu, \sigma^2) \Rightarrow z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

esto se llama estandarización.

Esto simplifica los cálculos pero no deja de ser complejo por eso se crea una tabla para los posibles resultados de la función de distribución acumulada.

$$\phi(z) = F_Z(z) = P(Z \leq z) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Si se desea calcular $P(a < X < b)$ entonces estandarizando tenemos:

$$P(a < X < b) = P(\frac{a-\mu}{\sigma} < \frac{x-\mu}{\sigma} < \frac{b-\mu}{\sigma})$$

$$= P(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}) = \phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma})$$