Capítulo 1: Introducción

Población, es el conjunto total de elementos sobre el cual uno va a realizar las mediciones individuales.
La población puede ser **FINITA** o **INFINITA**

Muestra, es un subconjunto de la población el cual se busca que sea representativo de la población (es decir, se comporte de manera acorde a la población).

Unidad Elemental, Es cada individuo de la población sobre el cual se va a medir la variable en estudio.

Variable, es cualquier característica que puede tomar diferentes valores en la unidad elemental.
Las variables pueden ser **CUALITATIVAS** o **CUANTITATIVAS**.

Las cualitativas son variables de tipo categóricas que definen categorías sin un orden en sí, por ejemplo “Color”, “Sexo”, “Nacionalidad”, etc. son variables categóricas.

Las cuantitativas se dividen a su vez en **DISCRETAS** o **CONTINUAS**.
Las variables discretas son aquellas las cuales poseen saltos numéricos, y por lo general usan los números naturales, por ejemplo la Edad en años, que se suele representar en 1,2,3,4,..., años.
Las variables continuas son aquellas que pueden adoptar cualquier valor real, un ejemplo podría ser la altura en de una persona.

Otra forma de agrupar las variables es según como funciona su escala de medición.

Nominal: Valores diferentes, que no poseen un orden. EJ: (Blanco, Azul, Rojo, Verde)
Ordinal: Valores diferentes que poseen un orden EJ: (Malo, Intermedio, Bueno)
De Razón: Posee un valor que se considera un 0 Absoluto y que tiene un significado en el problema. Un auto que tiene una cierta velocidad, su velocidad no puede poseer valores negativos, y si la velocidad vale 0, significa que este se encuentra quieto.
De Intervalo: No posee un 0 absoluto, el 0 es un valor que se puede asumir, pero no tiene un significado, por ejemplo la temperatura en grados Celsius, donde 0 ºC no significa la ausencia de calor.

En el caso de grados Kelvin, la temperatura es en escala de razón, ya que posee un 0 el cual implica la ausencia de calor.

Población estadística, son todos los valores los cuales puede asumir nuestra variable en estudio.

Parámetro, son los valores que resumen alguna información (característica) de la población en estudio.

Al plantear un problema siempre completaremos lo siguiente:

Población: …
Unidad Elemental: …
Variable/s en estudio: …
Población estadística: …
Parámetro de interés: …

Objetivo: … *(Puede ser analizar el comportamiento, es decir, su distribución u obtener un parámetro de interés para resumir una característica del sistema de estudio)*

Tipos de estudios estadísticos:

Observacional, es un estudio donde se medirán las características de la unidad elemental sin modificarla.

Por ejemplo, un estudio observacional podría ser, estudiar el color de pelo de personas dentro de un aula.

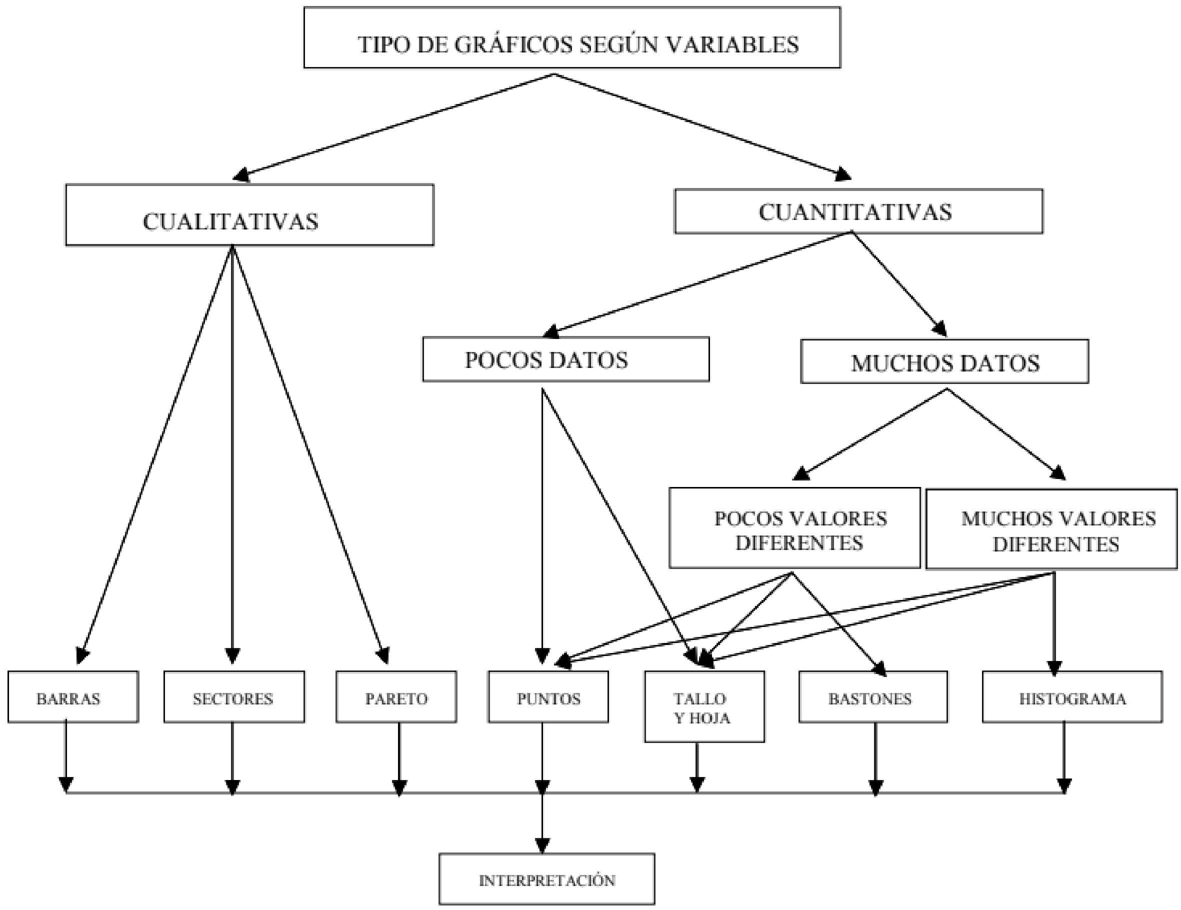
Experimental, es un estudio donde se busca cambiar variables del sistema para analizar una variable respuesta, busca responder a un causa-efecto.

Por ejemplo, para analizar la humedad de una habitación y uno podría cambiar la temperatura de la habitación, y luego medir la humedad.

Se puede estudiar toda la población, y obtener los parámetros a partir de la observación de todo el conjunto de población, (**CENSO**), o se puede aproximar el resultado a partir de la toma de una muestra que sea representativa de la población (**MUESTREO**) lo cual nos dará estimadores de los parámetros.

Capítulo 2: Estadística Descriptiva

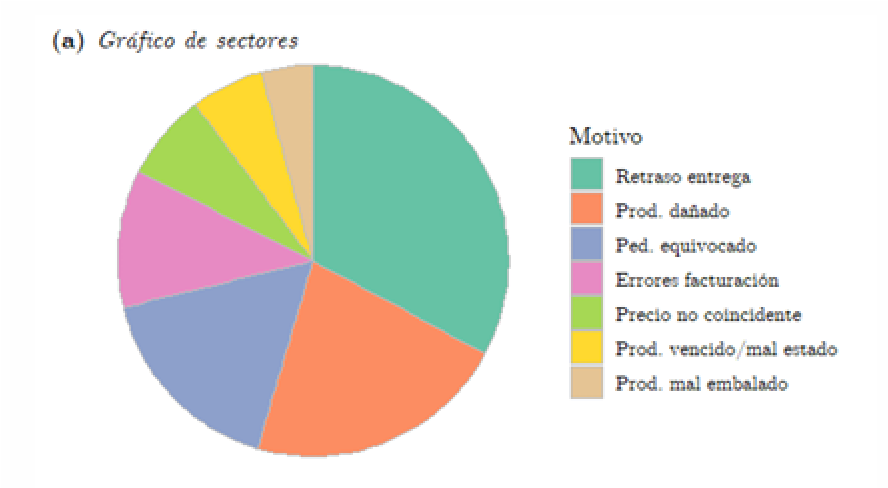
La analítica descriptiva posee distintos tipos de gráficos o formas de resumir información del sistema.

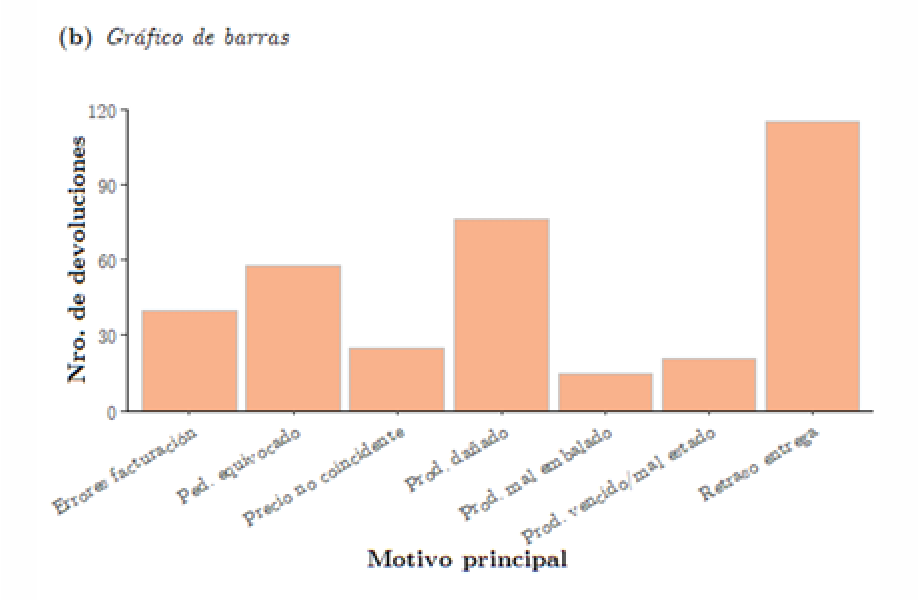


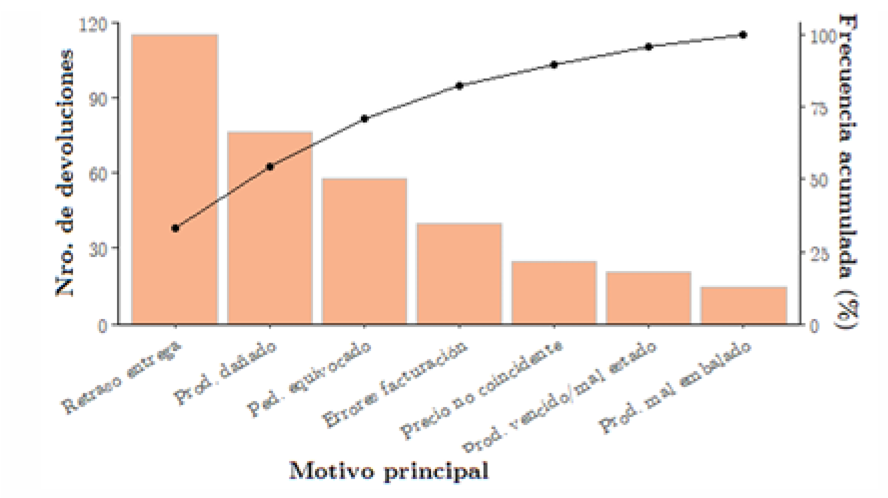
En primera instancia, uno puede poseer una tabla, en la tabla lo que uno puede realizar es una tabla de frecuencias.

Frecuencia Absoluta: Cantidad de veces que aparece un determinado valor. Supongamos en un caso de edades, la frecuencia absoluta de 34 años, sería la cantidad de individuos que presentaron 34 años.

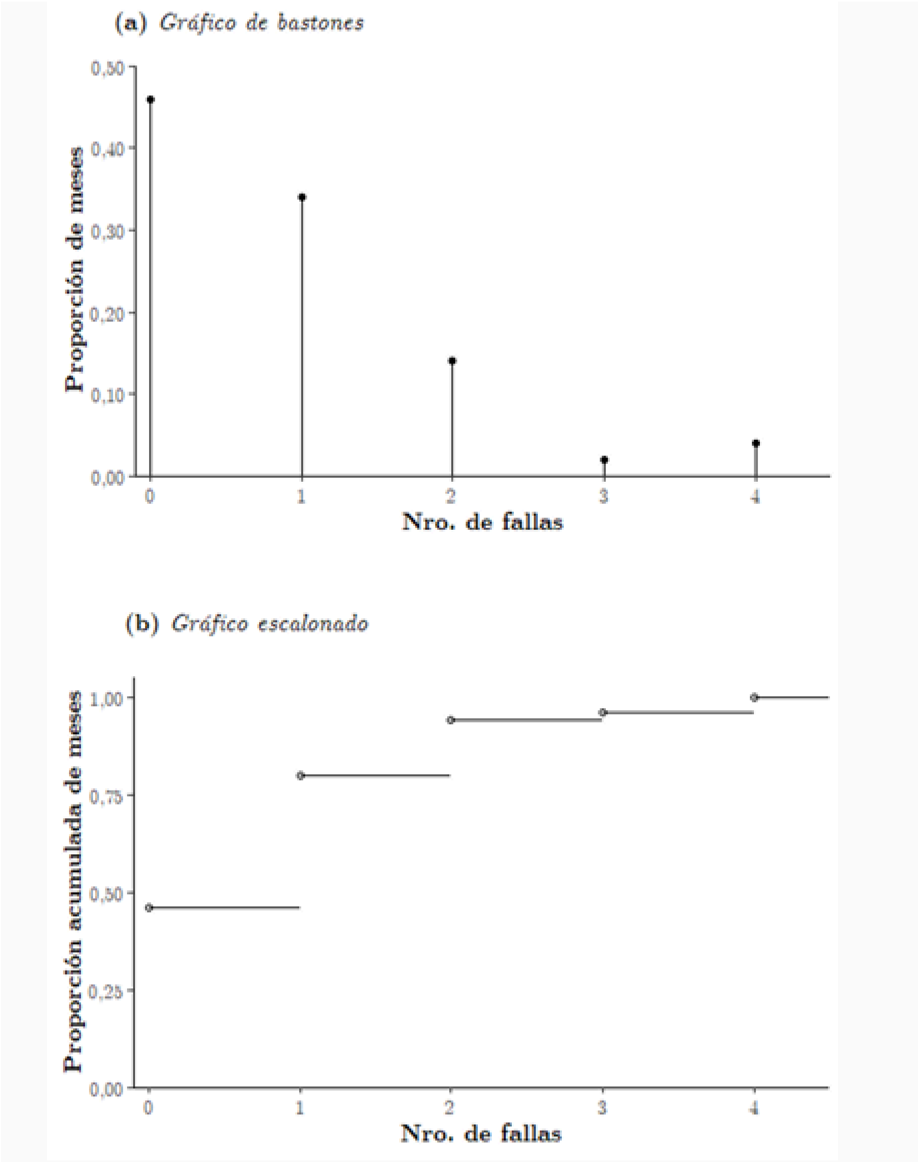
En los gráficos de sectores, barras y Pareto, están muy relacionados con la frecuencia. (Son graficos para variables **CUALITATIVAS**)



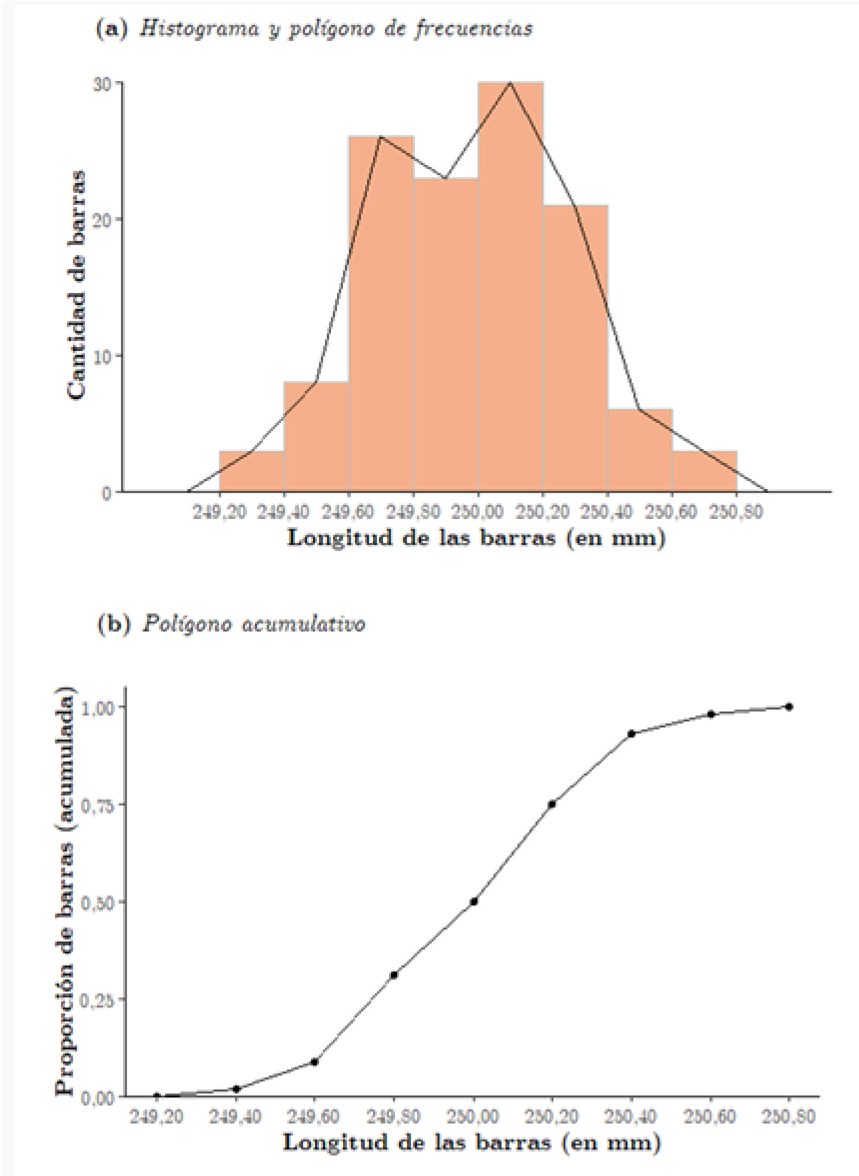




Los gráficos de bastones y escalonado, son para variables **CUANTITATIVAS DISCRETAS.**



Para las variables **CUANTITATIVAS CONTINUAS**, podemos utilizar intervalos para graficar un gráfico de barras, por ejemplo. O gráficos con polígonos de frecuencias, o polígono de frecuencia acumulada.



Otros tipos de diagramas para cuantitativas, son los siguientes, el de puntos, el de tallo y hoja:

Figura 2.13. *Diagrama de puntos para las longitudes de las barras analizadas.*

```
0 | 00000000000000000000000
1 | 00000000000000000
2 | 0000000
3 | 0
4 | 00
```

Figura 2.14. *Diagrama de tallo y hoja para el número de fallas por mes. La coma decimal se encuentra en la posición de la línea vertical.*

```
2492 | 9
2493 | 34
2494 | 59
2495 | 3688
2496 | 001222345588
2497 | 0133557777899
2498 | 0001125788
2499 | 002233334567789
2500 | 01334444555677888
2501 | 002444445688
2502 | 001234666699
2503 | 011334567
2504 | 0012
2505 | 068
2506 | 037
2507 | 1
```

Figura 2.15. *Diagrama de tallo y hoja para las longitudes de las barras analizadas. La coma decimal se encuentra una posición a la izquierda de la línea vertical.*

Medidas de resumen de los datos:

Se subdividen en dos grupos, de **POSICIÓN** y de **DISPERSIÓN.**

La posición nos da un resumen de sobre qué valor me encuentro, y las de dispersión de cuanto varía el valor que uno mide.

MEDIDAS DE POSICIÓN

$\mathbf{y_{min}}$ e $\mathbf{y_{max}}$**:** son los valores mínimo y máximo que toma la variable de estudio.

Percentiles: es el valor donde se acumula un porcentaje de los valores medidos, por ejemplo el percentil del 25% será un valor que los valores menores a él acumulen el 25% de todos los valores medidos.

Q1, primer cuartillo, es el percentil de 25%.

Mediana, es el segundo cuartillo cuyo percentil es el 50%.

Q3, es el tercer cuartillo el cual acumula un 75% de los datos.

Media aritmética (promedio): Se calcula como la suma de todos los valores, dividida el número total de valores.

Media truncada: Se calcula igual que la media aritmética, pero previo se ordena y se elimina un percentil (cota superior e inferior) de los datos. Permite eliminar valores atipicos.

Moda: Es el valor que se presenta con mayor cantidad de frecuencia absoluta.

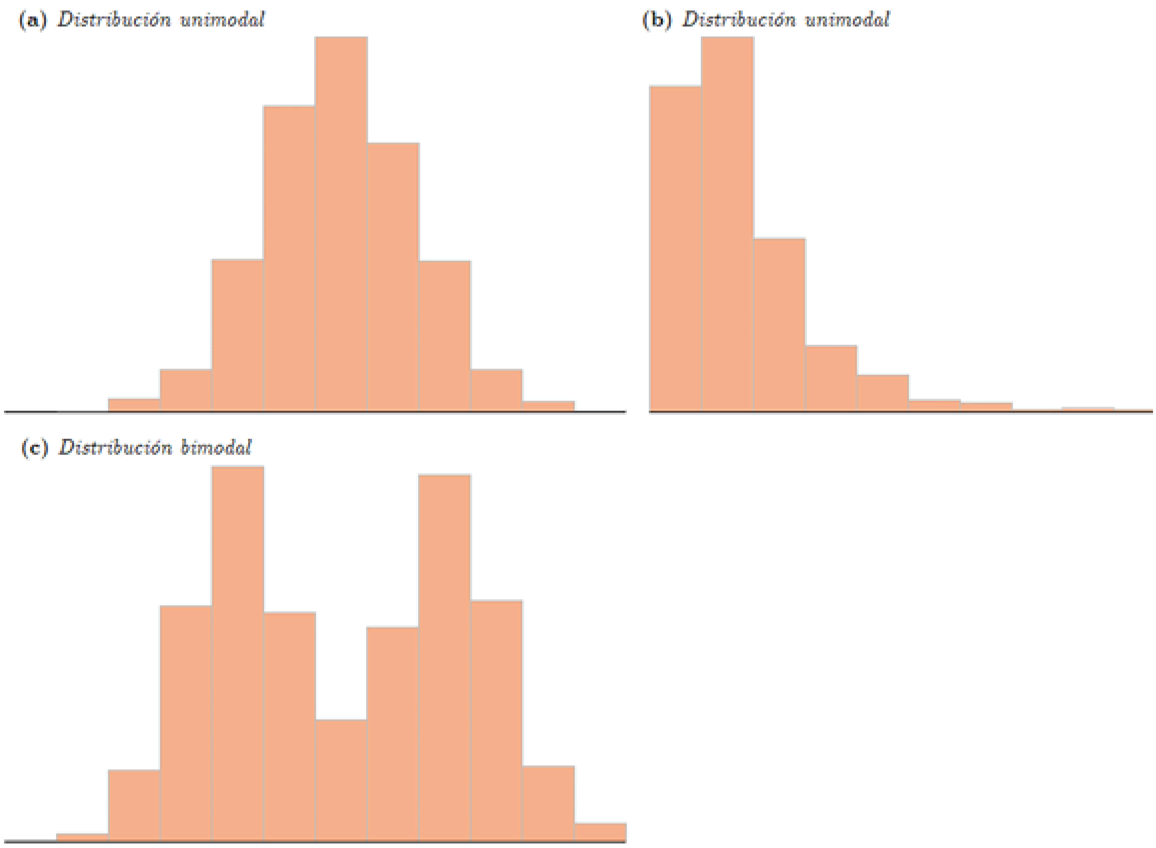


Figura 2.22. *Ejemplos de distribuciones de frecuencias unimodales y bimodales.*

MEDIDAS DE DISPERSIÓN

Rango: es la diferencia entre el y_{max} y el y_{min}

Rango intercuartil: es la diferencia entre q3 y q1.

Variancia: es el promedio de los desvíos cuadrados entre cada observación y el promedio.

Las expresiones que se presentan a continuación corresponden al cálculo de la varianza cuando los datos no están agrupados (2.8) y cuando están agrupados y en cada clase hay un único valor (2.9):

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}, \tag{2.8}$$

$$s^2 = \frac{\sum_{j=1}^{k}(y_j - \bar{y})^2 \cdot n_j}{n-1}, \tag{2.9}$$

donde $\bar{y}$ es el promedio o media aritmética, y_i es el valor i-ésimo del conjunto de datos, y_j es el valor correspondiente a la clase j, n es el total de datos observados, n_j es el número de datos observados para la clase j y k es la cantidad de clases.

Desviación estándar: es la raíz positiva de la variancia.

Coeficiente de variación: es el cociente entre la desviación estándar y el promedio multiplicado por 100. Expresa la desviación estándar como porcentaje del promedio.

La proporción es la frecuencia relativa asociada a una condición, la cantidad de veces que se da que se cumpla la condición sobre la cantidad total de medidas. f_0 en caso de muestra o π como parámetro poblacional.

Transformaciones lineales: si tenemos una variable X, la cual es una transformación lineal de una variable Y. Los estadísticos de X, están relacionados con los estadísticos de Y como lo indica la siguiente tabla.

Para las observaciones originales (y_i)	Para las observaciones transformadas ($x_i = a + by_i$)
$\bar{y}$	$\bar{x} = a + b \cdot \bar{y}$
s_y^2	$s_x^2 = b^2 \cdot s_y^2$
s_y	$s_x = \lvert b \rvert \cdot s_y$

El coeficiente de variación al ser una transformación lineal seguirá siendo el mismo.

Gráfico de boxplot: Se utiliza para obtener una buena cantidad de medidas resumen en un solo gráfico, y permite encontrar valores atípicos de manera rápida.

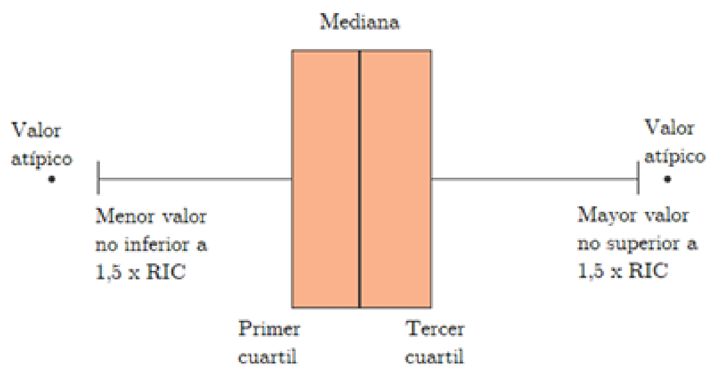
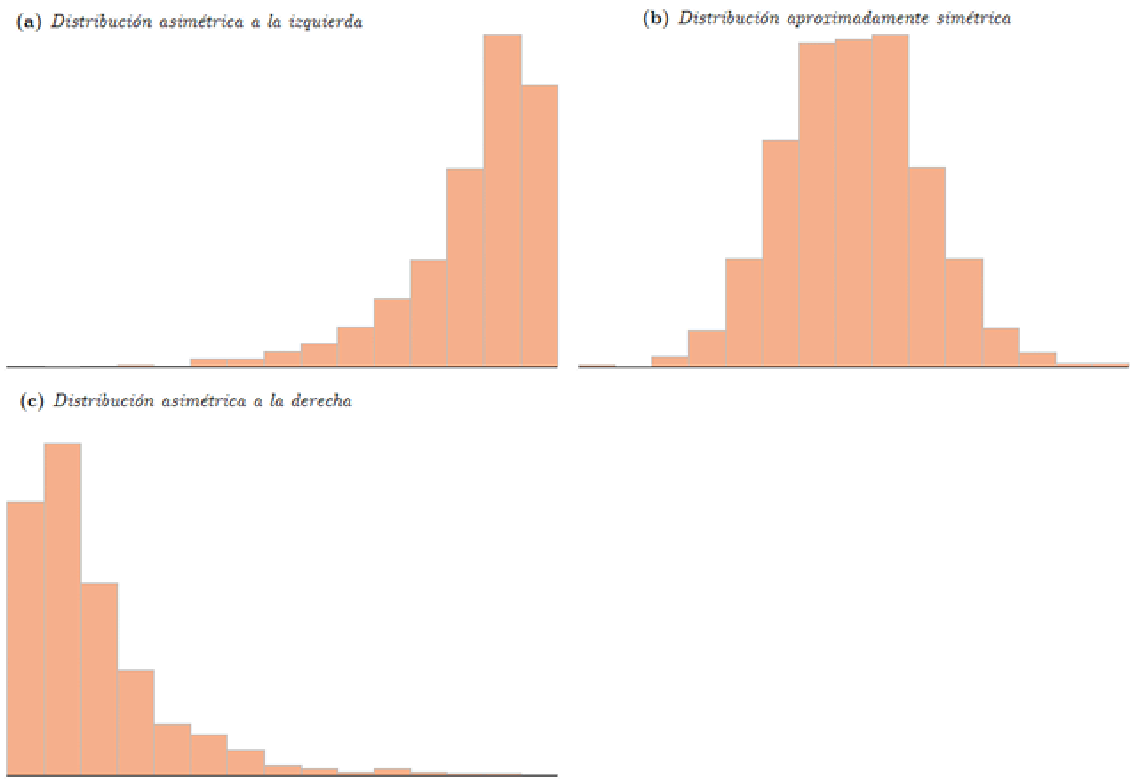


Figura 2.25. *Diagrama de Caja y Bigotes clásico utilizado para determinar si existen valores atípicos en un conjunto de datos.*

Otras medidas serán las de asimetría y aplastamiento.

Asimetría:



Aplastamiento o curtosis:

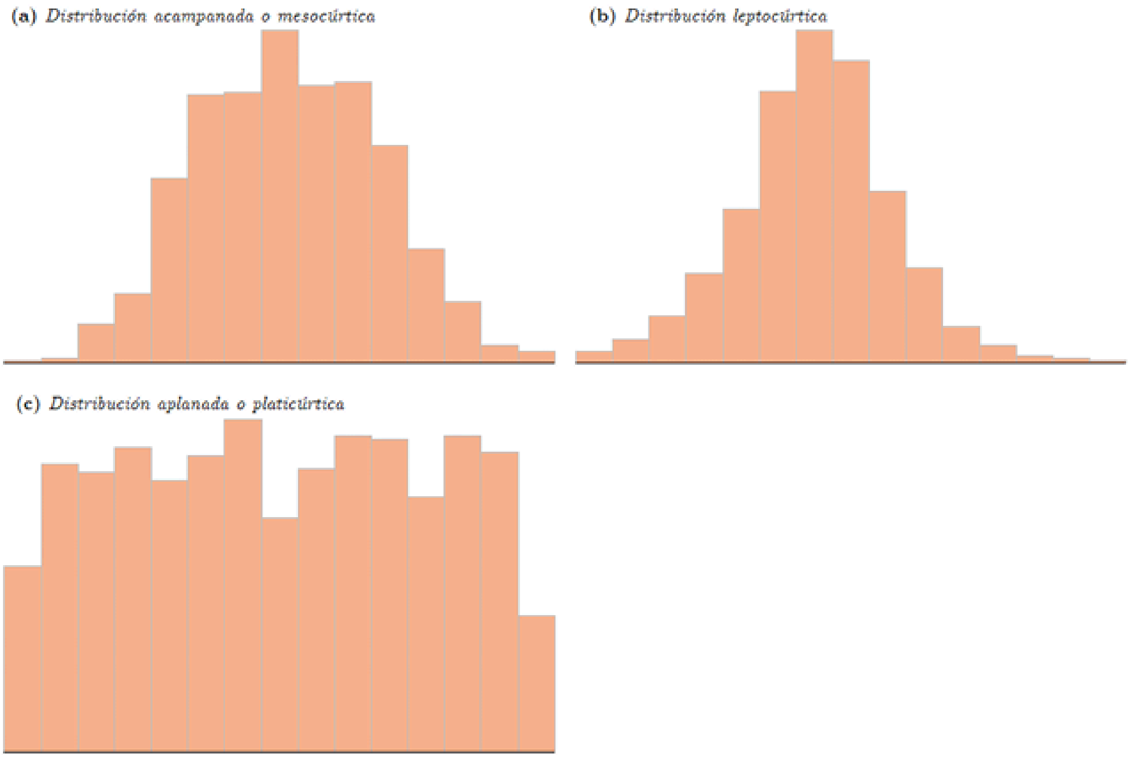


Figura 2.32. *Distribuciones con diferentes curtosis.*

Capítulo 3: Distribuciones de probabilidad

Variable aleatoria: Cada función Y que asocia un valor numérico a cada unidad elemental. (pueden ser CONTINUAS o DISCRETAS)

Función de densidad de probabilidad (Utilizada en variables aleatorias continuas):

Es una función continua la cual nos determina la probabilidad de que aparezca un valor.

- Está normalizada, es decir, su integral en todo su dominio es igual a 1, esto es coincidente con la definición de probabilidad de que varía entre 0 y 1.
- La probabilidad se calcula como el área bajo la curva de la f_y .
- El área de una línea es 0, por eso no se puede obtener una probabilidad puntual, salvo que se usen conceptos de límites.

- La función es siempre positiva o 0.

Función de probabilidad puntual (Utilizada en variables aleatorias discretas):

Es una función que va a obtener punto a punto la probabilidad de que se presente ese valor. El ejemplo más básico es tirar una moneda, donde podríamos definir cara o cruz como la variable 0 o 1, y la función será una que vale 0.5 si es cara y 0.5 si es cruz. p(0) = 0.5 y p(1) = 0.5

- Como se sabe, la sumatoria de las probabilidades debe ser igual a 1
- Además no existen probabilidades negativas, por eso debe valer entre 0 y 1 cualquier probabilidad puntual

Si tuviésemos infinitos puntos en esta función de probabilidad puntual, obtendremos una función de densidad.

- Acá no buscamos un área para definir la probabilidad, sino directamente el valor de la funcion en el punto.

Función de distribución acumulada:

Si quisiéramos obtener la probabilidad de que algo sea menor igual a un valor "x" vamos a plantear la P(≤x), esa es la funcion de distribucion acumulada.

Pensado para variable discreta va a ser la suma de todas las probabilidades de los valores menos a x. En variable continua es lo mismo, sin embargo, *como las probabilidades individuales son infinitesimales, tenemos que pensar que el infinito de pequeños forma un todo,* por eso necesitamos calcular el área bajo la curva desde -∞ a x. *Dato matemático: el concepto de sumatoria e integral son prácticamente lo mismo (justamente la integral es una suma de riemann, la diferencia es que la sumatoria sirve para un sistema discreto y la integral funciona para sistemas continuos)*

Función de distribución acumulada	Variable continua	Variable discreta
$F_Y(y)$	$\int_{-\infty}^{y} f_Y(s)ds$	$\sum_{s \leq y} p_Y(s)$

Esta función es importante, ya que es fácilmente trabajable. Supongamos que queremos obtener la probabilidad >x, como sabemos que la integral en todo el dominio es 1, es lo mismo que hacer 1 - P(≤x).
O supongamos que queremos obtener la probabilidad en un intervalo
P(a ≤ x ≤ b) = P(b) - P(a)

Medidas de resumen de una distribución de probabilidad:
Medidas de posicion
La media poblacional o la Esperanza, se puede calcular de la siguiente manera, como el producto entre la variable “y” y la función de probabilidad de “y”, $y \cdot f_Y(y)$, la diferencia entre discreta y continua será que en continua, esto se resolverá como una integral y en discreta como una sumatoria.
Este producto es fácil de entender si pensamos como calculamos un promedio. Un promedio es la suma de los valores dividido el total, “y” representa el valor, $f_Y(y)$ no es más que la cantidad de veces que se repite “y” sobre el total de valores.

$$\frac{1+2+1+4}{4} = \frac{1*2}{4} + \frac{2*1}{4} + \frac{4*1}{4} = 1\frac{2}{4} + 2\frac{1}{4} + 4\frac{1}{4}$$

Fíjense en el ejemplo anterior como convertimos el promedio “común” a un promedio en términos de la probabilidad de que aparezca cada valor.
Bueno, ese es discreto, y es una sumatoria del producto $y \cdot f_Y(y)$
Para continuo, entonces será una integral del producto $y \cdot f_Y(y)$

$$E(Y) = \mu_Y = \int_{y \in R_Y} y \cdot f_Y(y)dy;$$
$$E(Y) = \mu_Y = \sum_{y \in R_Y} y \cdot p_Y(y).$$

La mediana poblacional es aquel valor “y” que acumula una P(y) = 0.5
Entonces es fácil saber que directamente este será el f(y) = 0.5
En el caso de variable continua, será la integral hasta el punto que acumule 0.5 de la función de densidad, en caso de discreta será el valor hasta el cual la sumatoria acumule 0.5

$$F_Y(\tilde{\mu}_Y) = \int_{y \leq \tilde{\mu}_Y} f_Y(y) d_y = 0,5;$$
$$F_Y(\tilde{\mu}_Y) = \sum_{y \leq \tilde{\mu}_Y} p_Y(y) = 0,5.$$

Los cuartilos son de igual manera que la mediana, pero serán 0.25 el primer cuartillo y 0.75 el tercer cuartilo.

La moda es nuevamente el valor "y" que posee el mayor valor de probabilidad, es decir, el valor más alto de la función de densidad o el mayor valor de probabilidad en caso de ser discreto.

Medidas de dispersión

El desvío estándar poblacional, nuevamente por definición, en continuo y discreto, en uno es una sumatoria y en otro es una integral.

Y la definición es la misma, es el cuadrado de la diferencia (y - μ_y) eso es cuanto se desvía un valor con respecto a la media poblacional o esperanza.

Todo eso multiplicado por… La probabilidad

El **desvío estándar poblacional** de la variable aleatoria Y se simboliza con $D(Y)$ o σ_Y.

Informalmente, se lo define como la raíz cuadrada de la esperanza matemática de los desvíos con respecto a la media poblacional elevados al cuadrado, es decir:

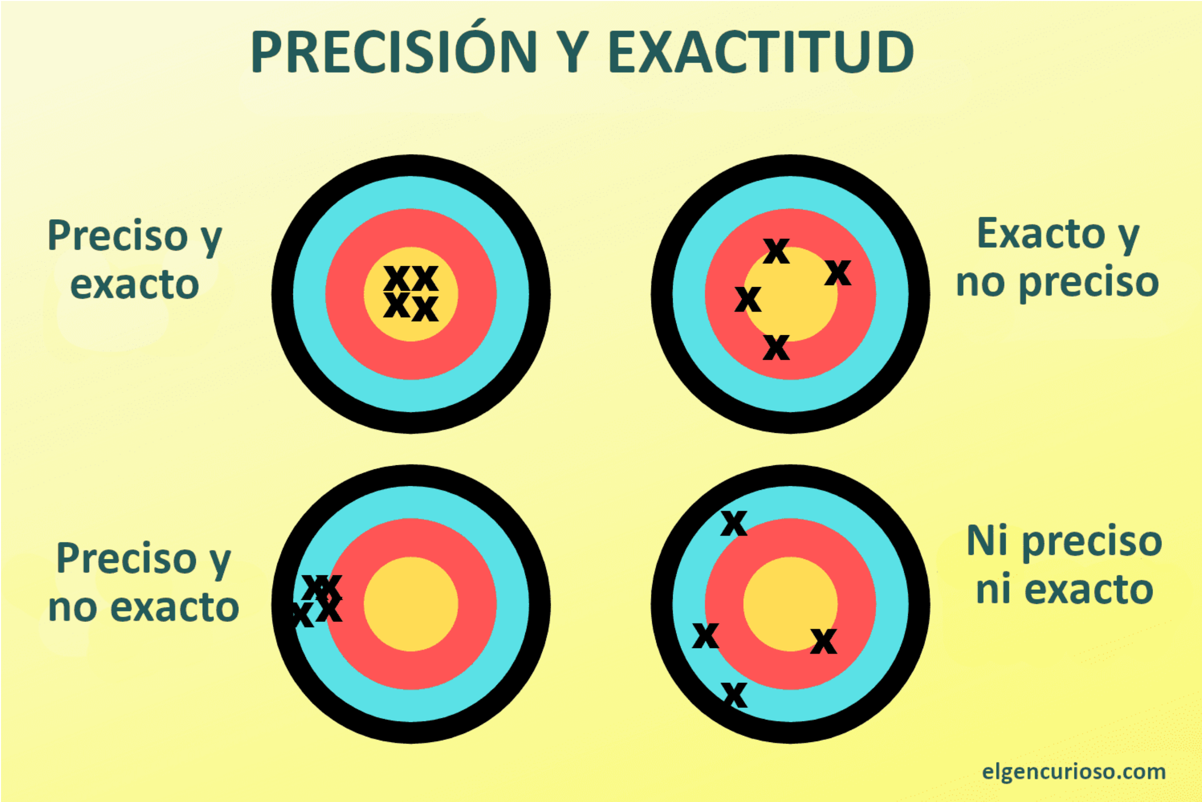
- cuando Y es continua, $D(Y) = \sigma_Y = \sqrt{E(Y-\mu_Y)^2} = \sqrt{\int_{y \in R_Y} (y-\mu_Y)^2 . f_Y(y) d_y} =$
$= \sqrt{\int_{y \in R_Y} y^2 \cdot f_Y(y) d_y - \mu_Y^2} = \sqrt{E(Y^2) - \mu_Y^2};$
- cuando Y es discreta, $D(Y) = \sigma_Y = \sqrt{E(Y-\mu_Y)^2} = \sqrt{\sum_{y \in R_Y} (y-\mu_Y)^2 \cdot p_Y(y)} =$
$= \sqrt{\sum_{y \in R_Y} y^2 \cdot p_Y(y) - \mu_Y^2} = \sqrt{E(Y^2) - \mu_Y^2}.$

El rango intercuartílico es la diferencia entre el Q3 y el Q1

La medición: A la hora de medir una variable existen dos conceptos clave, la exactitud y la precisión.

Exactitud: El promedio de la medición se condice con el promedio general

Precisión: Cada medición que yo realizo varía muy poco con respecto a la anterior.



Preciso y exacto, el promedio está en el centro, y los puntos varían muy poco (arriba a la izquierda)

Exacto e impreciso, el promedio está en el centro, pero los resultados varían mucho.

Preciso e inexacto, el promedio NO está en el centro, pero los resultados varían poco.

Impreciso e inexacto, el promedio NO está en el centro, y los resultados varían mucho.

Expresado matemáticamente, el valor medido X, sera el valor real δ + un error constante (sesgo) + un error aleatorio

$X = \delta + k + \varepsilon$

El sesgo k, es un error constante, el cual nos habla de la exactitud, por ejemplo siempre medimos por demás. Esto se ve reflejado en la esperanza, si está sesgado la $E(X) = \delta + k$

En cambio, si no hay sesgo o error sistemático entonces $E(X) = \delta$

porque k valdría 0.

El error aleatorio, como puede variar erráticamente, está asociado a la variación y no a la esperanza. Supongamos el caso perfecto donde no existe error aleatorio, siempre ni sesgo, siempre que tomáramos una medida obtendremos X = δ, por lo cual no habría variabilidad. Entonces la variabilidad, es decir, la precisión estará dada por este error aleatorio.

$$V(X) = V(\delta + k + \varepsilon) = V(\varepsilon)$$

La variancia de δ y de k son 0, ya que son valores constantes.

La precisión entonces se determina a partir de fijar un valor tolerable de variación que dependerá del problema.

Síntesis:

	Variable continua	Variable discreta
Distribución de probabilidad	f_Y: función de densidad de probabilidad	p_Y: función de probabilidad puntual
Probabilidades	$P(Y = y_1) = 0$ $P(Y < y_1) = \int_{-\infty}^{y_1} f_Y(y)dy$ $P(Y > y_1) = \int_{y_1}^{+\infty} f_Y(y)dy$ $P(y_1 \leq Y \leq y_2) = \int_{y_1}^{y_2} f_Y(y)dy$	$P(Y = y_1) = p_Y(y_1)$ $P(Y < y_1) = \sum_{y<y_1} p_Y(y)$ $P(Y > y_1) = \sum_{y>y_1} p_Y(y)$ $P(y_1 \leq Y \leq y_2) = \sum_{y=y_1}^{y_2} p_Y(y)$
$F_Y(y)$	$\int_{-\infty}^{y} f_Y(s)ds$	$\sum_{s\leq y} p_Y(s)$
$E(Y) = \mu_Y$	$\int_{y\in R_Y} y.f_Y(y)dy$	$\sum_{y\in R_Y} y.p_Y(y)$
$D(Y) = \sigma_Y$	$\sqrt{\int_{y\in R_Y} y^2.f_Y(y)d_y - \mu_Y^2}$	$\sqrt{\sum_{y\in R_Y} y^2.p_Y(y) - \mu_Y^2}$

Capítulo 4: Probabilidad

Experimento aleatorio, se selecciona una unidad al azar, y se mide la variable de interés.

Pensemos en un dado, contar la cantidad de caras del dado puede ser determinista si considero que tengo todos dados de 6 caras. No sería aleatorio (Determinista), en cambio, tomar un dado y tirarlo y ver qué cara sale, es No determinista, porque no sé qué valor saldrá cada vez que lo tire. Para un experimento aleatorio válido, se debe establecer el conjunto de posibles resultados (cada valor de cada cara) y se puede repetir gran cantidad de veces en iguales condiciones (tirar muchas veces el dado).

Espacio muestral: Conjunto de resultados de un experimento aleatorio.
Resultado o punto muestral: Cada elemento del espacio muestral, todo resultado siempre se puede obtener al llevar a cabo el experimento y estos resultados son mutuamente excluyentes entre si, si sucede uno no puede suceder el otro.
Sucesos o eventos: Son subconjuntos del espacio muestral.

Si se define una variable aleatoria Y asociada a un experimento aleatorio dado, el espacio muestral es el **recorrido**, R_Y, de dicha variable y los sucesos pueden expresarse en función de Y. Por ejemplo, $M = \{Y : Y \leq 2\}$, $N = \{Y : 1 \leq Y < 8\}$, etc.

Los sucesos de un espacio muestral pueden tener operaciones de conjuntos entre si. (llamemos a los sucesos A y B)
El **complemento** de A es Ā, el complemento es el subconjunto que hace que S = A + Ā, es decir, que todos los elementos de S que no pertenezcan a A pertenecen a Ā
La **unión,** son los sucesos que pertenecen a A o a B
La **intersección**, son los sucesos que pertenecen a A y a B al mismo tiempo

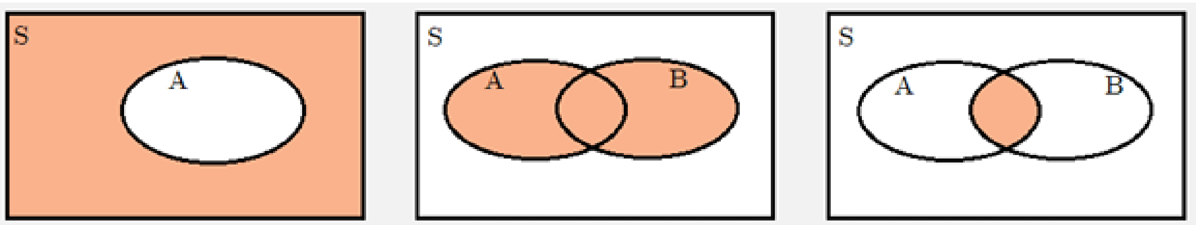


Figura 4.1. *Representación de los sucesos $\bar{A}$, $A \cup B$ y $A \cap B$, respectivamente.*

Mutuamente excluyentes, la intersección de A y B es vacía.

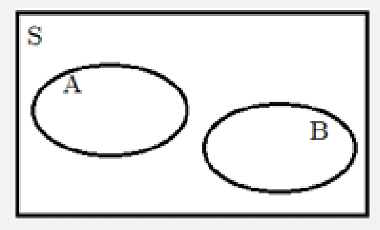


Figura 4.2. *Representación del suceso imposible como la intersección vacía entre A y B. En símbolos:* $A \cap B = \emptyset$.

Probabilidad: La probabilidad es la posibilidad de que suceda un suceso, es decir, sobre la cantidad de experimentos que realizo, qué cantidad de sucesos favorables sucedieron.

Caso 1:

En casos como el del DADO, que todas las caras son X probables, tenemos un caso como el siguiente.

Si en el espacio muestral existen k sucesos elementales equiprobables, la probabilidad de cada uno de ellos será entonces $\frac{1}{k}$, para asegurar que la suma total resulte 1. Luego, la probabilidad de un suceso compuesto A (que contiene c sucesos elementales) será igual a $\frac{c}{k}$, lo que da lugar a la regla:

$$P(A) = \frac{\text{número de resultados (o casos) favorables al suceso A}}{\text{número de casos o resultados posibles del experimento aleatorio}} = \frac{c}{k}.$$

Caso 2:

Si conocemos la función de probabilidad (densidad o puntual) podemos obtener también las probabilidades de los sucesos, a partir de la probabilidad acumulada.

Caso 3:

Denotando con n al número de repeticiones de la experiencia, y con $n(A)$ al número de veces que ocurre el evento A entre esas n repeticiones, resulta:

$$P(A) = lim_{n \to \infty} \frac{n(A)}{n} = lim_{n \to \infty} f(A).$$

Esto es un caso generalizado del caso 1, la diferencia entre el caso 1 y el caso 3 es el concepto de A priori y a Posteriori.

En el caso 1, nosotros de partida ya podemos predecir la probabilidad, en un dado de 6 caras sabemos que la probabilidad de cada cara es 1 sexto. Los resultados posibles son 6 y cada cara puede salir 1 sola vez, entonces con el caso 1 nosotros podríamos decir que la probabilidad de que salga 6 es ⅙.

Ahora supongamos que queremos ver si un dado está funcionando mal (está cargado), deberíamos tirar el dado un montón de veces, y contabilizar la cantidad de veces que salga 6, dividirlo sobre el número de veces que tiramos el dado. Supongamos que tiramos 100 veces el dado y sale 60 veces el número 6. La P(6) = 60/100 = 0.6 . Entonces, como es mucho mayor al ⅙ que esperábamos como resultado, podríamos afirmar que el dado está cargado. Pero alguien podría plantear que tirar 100 veces es poco, entonces ahí es donde en el caso 3 existe el límite tendiendo a infinito, ese número de repeticiones tiene que ser muy grande.

Reglas para la Probabilidad:

Axiomas de la probabilidad de un suceso

1. $P(A) \geq 0$ para cualquier suceso A,
2. $P(S) = 1$,
3. si A y B son sucesos mutuamente excluyentes y se considera la unión de ellos llamando C al nuevo suceso (es decir $C = A \cup B$), resulta

$$P(C) = P(A \cup B) = P(A) + P(B).$$

$P(A) = 1 - P(\bar{A})$

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

Esa resta es porque la intersección se estaría contando 2 veces, por eso la restamos. Si son mutuamente excluyentes, esa intersección vale 0.

La probabilidad condicional

La probabilidad de un suceso A condicionada a otro suceso B del mismo espacio muestral, denotada por $P(A/B)$ y denominada **probabilidad condicional**, es la probabilidad del suceso A condicionada a la ocurrencia del suceso B (es decir, en un espacio muestral restringido por el suceso B):

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \text{ para } P(B) \neq 0.$$

Sucesos independientes:

Se dice que A y B son **independientes** si la ocurrencia de uno de ellos no modifica las probabilidades de ocurrencia del otro.

Es decir, si $P(A/B) = P(A)$ (ó $P(B|A) = P(B)$, ya que si se da una de estas igualdades, se verifica también la otra).

Cuando se verifica la independencia entre A y B, resulta $P(A \cap B) = P(A) \cdot P(B)$.

A modo de resumen:

- si A y B son sucesos mutuamente excluyentes, resulta $P(A \cup B) = P(A) + P(B)$;
- $P(A) = 1 - P(\bar{A})$;
- si A y B son sucesos cualesquiera, resulta $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
- $P(A/B) = \frac{P(A \cap B)}{P(B)}$, para $P(B) \neq 0$;
- $P(A \cap B) = P(A/B).P(B)$, para $P(B) \neq 0$;
- si A y B son independientes, resulta $P(A \cap B) = P(A) \cdot P(B)$.

DIAGRAMA DE ÁRBOL:

Fijarse que dividimos primero en la probabilidad de que suceda L1 y L2, cuya probabilidad es de 0,5. Ahora sobre L1 vemos cuál es la probabilidad de que sean buenas, regulares o malas (esto nos dan 3 sucesos que su suma debería dar 1 en términos de Probabilidad)

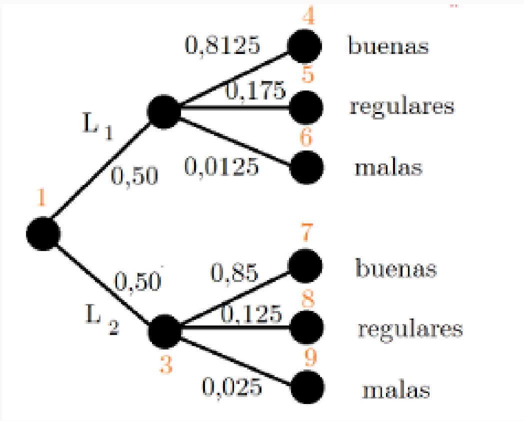


Figura 4.5. *Esquema de árbol.*

El primer conjunto de ramas se refiere al lote de donde se extrae la placa: en cada rama se representa uno de los lotes (suceso L_i) y su probabilidad ($P(L_i)$), para $i = 1,2$.

La probabilidad de que la placa seleccionada provenga del Lote 1 vale 0,50 ($P(L_1 = \frac{80}{160} = 0,50$).

La probabilidad de que la placa seleccionada provenga del Lote 2 vale 0,50 ($P(L_2 = \frac{80}{160} = 0,50$).

Observe que los sucesos L_1 y L_2 son mutuamente excluyentes y sus probabilidades suman 1.

Capítulo 5: Distribuciones de Probabilidad de uso frecuente

Distribución Normal (o Campana de Gauss): Tiene forma de campana.
Es la función más común y clave en la estadística.
Viene definida principalmente por la media poblacional y la desviación estándar poblacional.

Una variable aleatoria continua Y tiene una distribución Normal de parámetros matemáticos[a] μ y σ, y se simboliza $Y \sim N(\mu;\sigma)$, si su función de densidad de probabilidad es:

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \ con\ y \in \mathbb{R}$$

donde $\mu \in \mathbb{R}$ y $\sigma \in \mathbb{R}^+$.

[a] Son aquellos valores, que si se conocen, hacen que la función quede completamente definida.

Tiene forma de campana, (es simétrica para ambos lados)

En el medio tenemos la media, la cual a su vez también va a ser la mediana y la moda.

El valor que tiene en el promedio (μ) es $\frac{1}{\sqrt{2\pi\sigma}}$, el cual es el máximo de la función.

Si variamos μ se mueve la campana, si variamos σ se aplana o se estira.

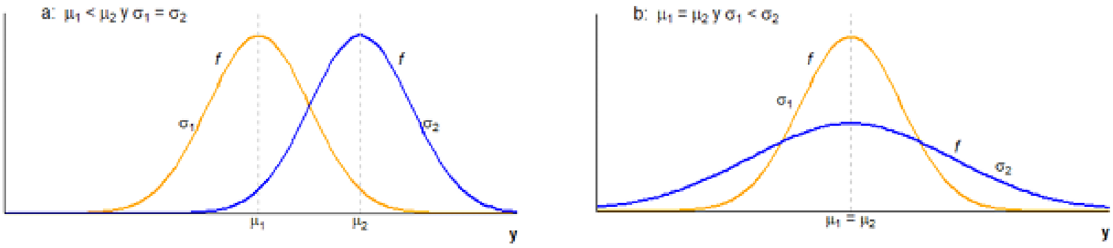


Figura 5.2. *Cambios en la curva Normal al variar μ o σ.*

Su función de probabilidad acumulada es:

La función de distribución acumulada Normal está definida como:

$$F_Y(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} ds, \; con \; y \in \mathbb{R}.$$

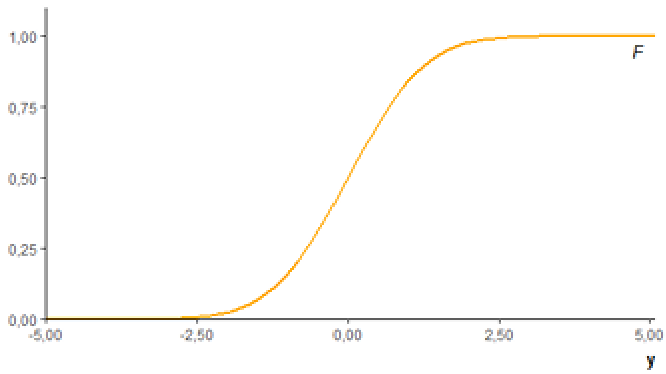


Figura 5.3. *Función de distribución acumulada Normal.*

Muy importante es saber que, La esperanza de una variable con distribucion normal es igual a su media aritmética, y su desviación estándar es la desviación estándar poblacional.

Si $Y \sim N(\mu;\sigma)$, se demuestra que su media es $E(Y) = \mu$ y su desvío estándar es $D(Y) = \sigma$.

Nuevamente, para obtener una Probabilidad de una variable con distribución normal, basta con calcular el área bajo la curva de la campana de gauss hasta el valor que uno desea. (Como esto puede resultar “Complicado” de resolver el área analíticamente, “Jorge Gauss” un tipito simpático, armo tablas donde si uno normaliza una distribución Gaussiana con la definición de Z, puede obtener la probabilidad.

(Es decir, definimos $z = \frac{Y-\mu}{\sigma}$) esto nos da una gaussiana que… tiene μ_z = 0 (promedio en 0) y desviacion estandar σ_z = 1.

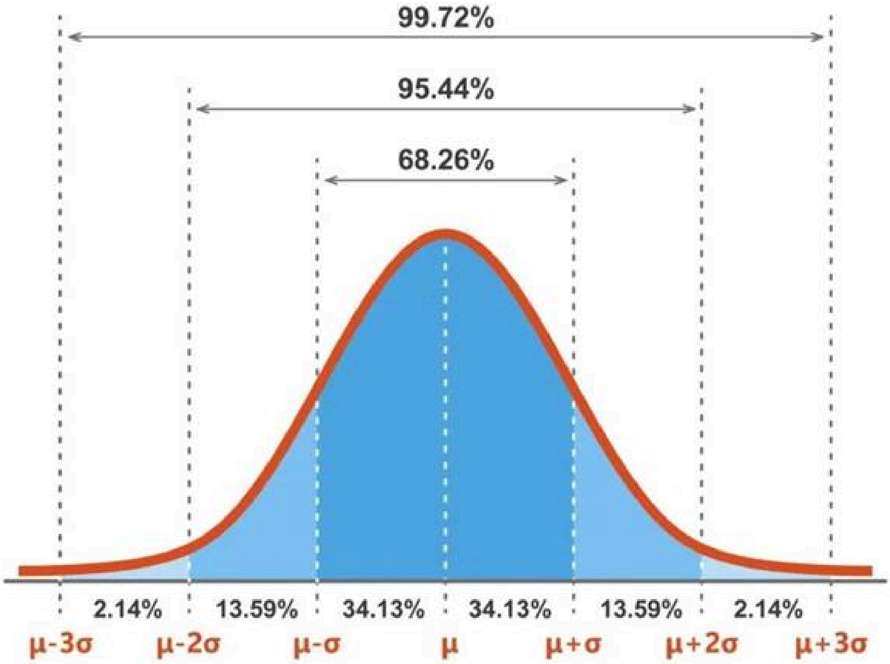
“”

Entonces para obtener una probabilidad calculamos z, buscamos el valor de z en la tabla y vemos que probabilidad tiene.

Ahora, si con una probabilidad queremos el valor Y, vamos a la tabla, buscamos la probabilidad que más o menos de, revisamos el z que corresponde a esa probabilidad y despejamos Y de la definición de z
“”

Si no, la forma para los Pillos, es la siguiente:
Conocer las probabilidades de acuerdo al promedio ± $n\sigma$.
La probabilidad en los rangos es:
$\mu \pm \sigma$ = 68.2 %
$\mu \pm 2\sigma$ = 95.44 %
$\mu \pm 3\sigma$ = 99.72 %
Se ve clarito en el grafico, que como es simetrica (y teniendo en cuenta que hasta el promedio acumulamos 50% de probabilidad, podemos jugar con eso

para resolver casi cualquier problema o en todo caso “aproximarlo”.



Distribución uniforme:

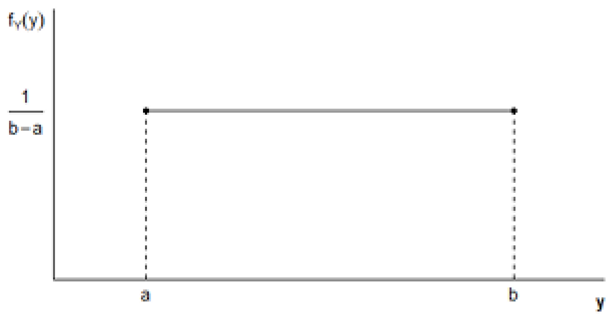


Figura 5.8. *Función de densidad de probabilidad Uniforme de parámetros a y b.*

Una variable aleatoria continua Y tiene una distribución Uniforme de parámetros matemáticos a y b, y se simboliza $Y \sim U(a,b)$, si su función de densidad de probabilidad es:

$$f_Y(y) = \begin{cases} \frac{1}{b-a} \; si \; a \leq y \leq b \\ \\ 0 \; si \; y < a \; o \; y > b \end{cases}$$

donde a y $b \in \mathbb{R}$.

Si $Y \sim U(a;b)$, se demuestra que su media es $E(Y) = \frac{a+b}{2}$ y su desvío estándar es $D(Y) = \sqrt{\frac{(b-a)^2}{12}}$.

Calcular probabilidad es fácil, área bajo la curva, es una multiplicación básica (base por altura).

La esperanza fíjense que es el promedio a fin de cuentas, no más que como es constante haces los extremos divido 2 y te da.

Y el desvío estándar sale de definición.

Distribución triangular:

Una variable aleatoria continua Y tiene una distribución Triangular de parámetros matemáticos a, b y c, y se simboliza $Y \sim Tri(a;b;c)$, si su función de densidad de probabilidad es:

$$f_Y(y) = \begin{cases} \frac{2(y-a)}{(b-a)(c-a)} & si \quad a \leq y < c \\ \frac{2}{b-a} & si \quad y = c \\ \frac{2(b-y)}{(b-a)(b-c)} & si \quad c < y \leq b \\ 0 & si \quad y \notin [a,b] \end{cases}$$

donde $a,\ b,\ c \in \mathbb{R}$

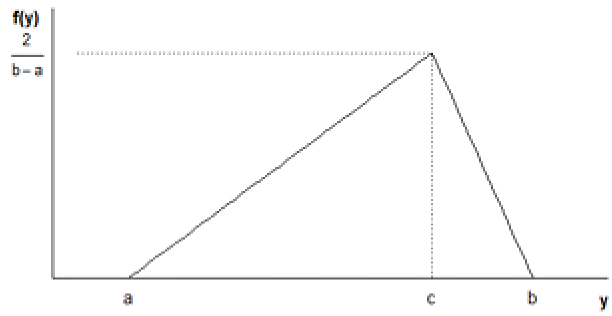


Figura 5.12. *Función de densidad de probabilidad Triangular de parámetros a, b y c.*

Si $Y \sim Tri(a;b;c)$, se demuestra que su media es $E(Y) = \frac{a+b+c}{3}$ y su desvío estándar es $D(Y) = \sqrt{\frac{a^2+b^2+c^2-ab-ac-bc}{18}}$.

Si nos fijamos la esperanza es lo mismo que antes, a + b + c / 3 , osea un promedio y la desviacion estandar nuevamente sale por definicion, (esta es mas quilombo) pero tanto esta como la anterior es la raiz de una diferencia que debe dar positiva / n * 6
antes era 2 * 6 = 12 , siendo 2 a y b, ahora con c pasa a ser 3 * 6 = 18.
Si agregáramos un d, sería 4 * 6 = 24
(ESTO LO INFIERO, pero va por ahi seguro)

Distribución Exponencial:
Se define por su parámetro ALPHA α

Una variable aleatoria continua Y tiene una distribución Exponencial de parámetro matemático α, y se simboliza $Y \sim Exp(\alpha)$, si su función de densidad de probabilidad es:

$$f_Y(y) = \begin{cases} \alpha e^{-\alpha y} & si \quad y \geq 0 \\ 0 & si \quad y < 0. \end{cases}$$

donde $\alpha > 0$

Solo tiene sentido verla de 0 en adelante, es decreciente, asimetrica a la derecha. y la mayor probabilidad esta en 0.
La media, mediana y moda son todas distintas.

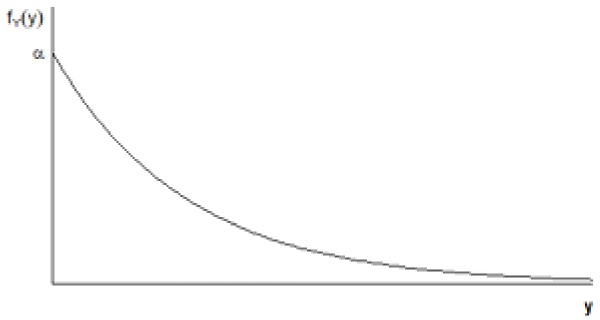


Figura 5.16. *Función de densidad de probabilidad Exponencial de parámetro α.*

Nuevamente, para obtener un valor deberíamos hacer el área bajo la curva. Acá si es interesante la función de probabilidad acumulada.

La función de distribución acumulada Exponencial está definida como:

$$F_Y(y) = \begin{cases} 0 & si \quad y < 0 \\ 1 - e^{-\alpha y} & si \quad y \geq 0 \end{cases}$$

Básicamente, reemplazamos y en la función para los y mayores iguales que 0 y listo, nos da la probabilidad.

Si $Y \sim Exp(\alpha)$, se demuestra que su media es $E(Y) = \frac{1}{\alpha}$ y su desvío estándar es $D(Y) = \frac{1}{\alpha}$.

Distribuciones Discretas:

Distribucion binomial: Esta es una distribución donde se genera un árbol de probabilidad, entre un suceso A y su suceso contrario Ā. Donde las probabilidades serán π y 1 - π.

El "ancho" del árbol vendrá representado por el n, que es la cantidad de veces que repito.

Supongamos tirar una moneda n veces, π = 0.5 y 1-π = 0.5, y n será el número de veces que tire yo la moneda.

En el árbol graficado abajo sería, Cara o Cruz la primera, después si saque Cara, la segunda vez que tire, que salga Cara o Cruz, y así sucesivamente.

En la columna que tiene enteros, nos dice la cantidad de veces que salió cara.

Y al final la forma de sacar la probabilidad de esa conformación en particular.

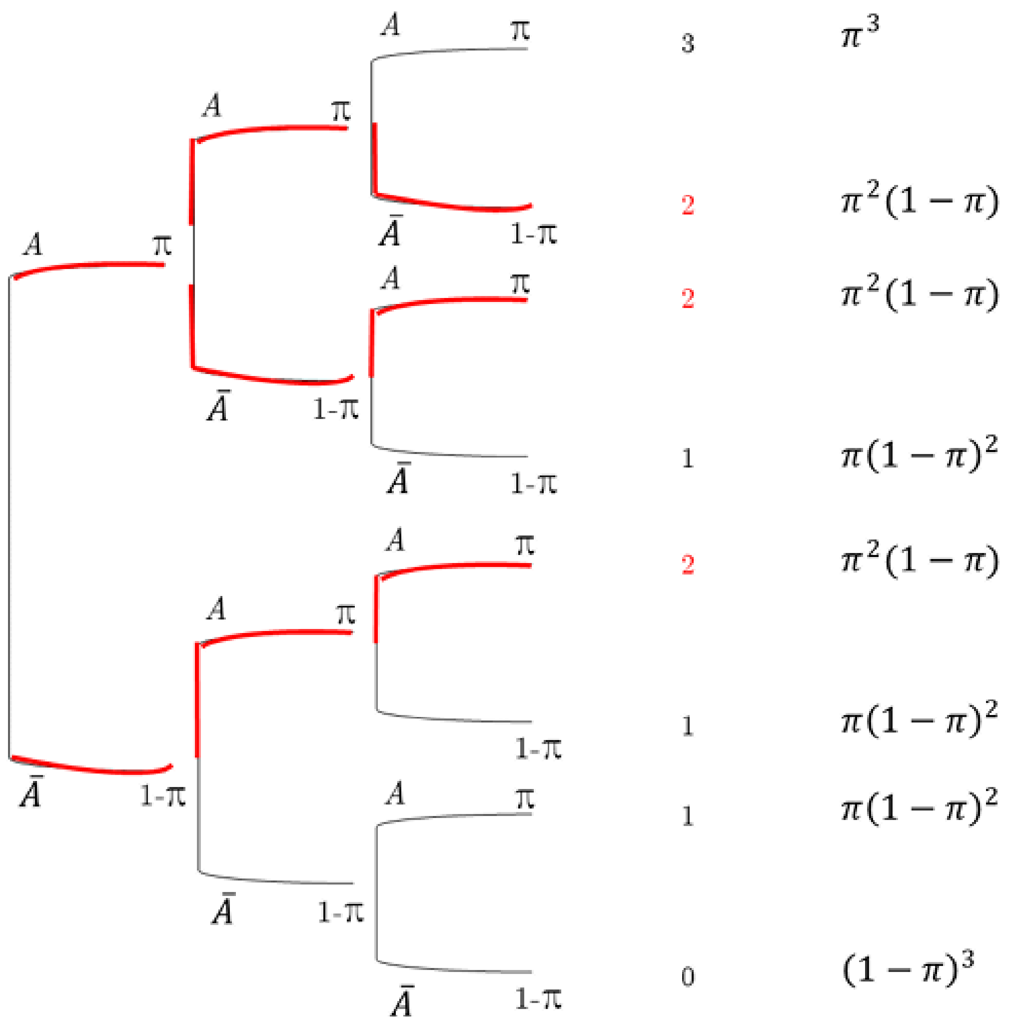


Figura 5.22. *Árbol de probabilidades para una distribución Binomial de parámetros n=3 y π.*

Generalizamos como un producto de $\pi^A (1 - \pi)^{n - A}$

Supongamos que en el árbol queremos sacar todas las posibilidades de que suceda 2 veces A, entonces M = 2, analizamos en el diagrama de arbol cuantas veces se repite el M = 2, y vemos que 3 veces.

Entonces $P(2) = 3 * \pi^2 (1 - \pi)^1$

Una variable aleatoria discreta Y tiene una distribución Binomial con parámetros matemáticos n y π, y se simboliza $Y \sim Bi(n;\pi)$, si su función de probabilidad puntual es:

$$p_Y(y) = \binom{n}{y} \pi^y.(1-\pi)^{(n-y)} \; con \; y = 0, 1, \ldots, n \; ^a$$

donde $n \in \mathbb{N}$ y $\pi \in [0, 1]$.

[a] La expresión hace referencia a las "combinaciones de n elementos tomadas de y" es decir, al número de grupos distintos (si difieren de un elemento sin importar el orden) de tamaño y que se pueden formar a partir de un total de n elementos. Se obtiene de la siguiente manera: $\binom{n}{y} = \frac{n!}{y!(n-y)!}$

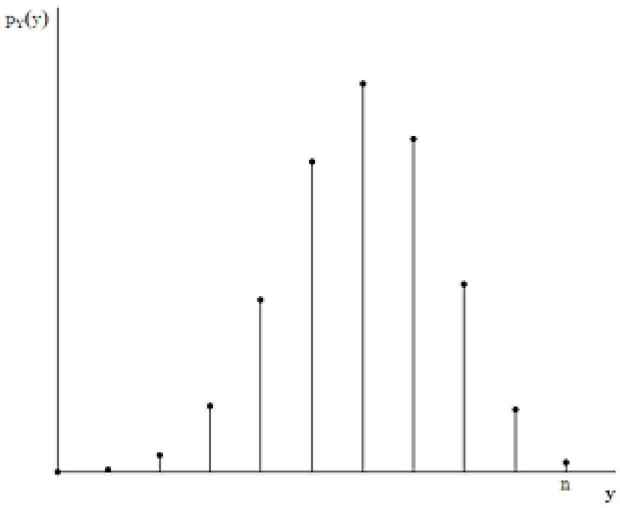


Figura 5.23. *Función de probabilidad puntual Binomial de parámetros n y π.*

Si $Y \sim Bi(n,\pi)$, se demuestra que su media es $E(Y) = n \cdot \pi$ y su desvío estándar es $D(Y) = \sqrt{n \cdot \pi \cdot (1-\pi)}$.

Distribución de Poisson:

La distribución de Poisson viene determinada por un solo parámetro α

Una variable aleatoria discreta Y tiene una distribución Poisson con parámetro matemático α (con $\alpha \geq 0$), y se simboliza $Y \sim Po(\alpha)$, si su función de probabilidad puntual es:

$$p_Y(y) = \frac{e^{-\alpha}\alpha^y}{y!} \; con \; y = 0, \; 1, \; 2, \; \ldots$$

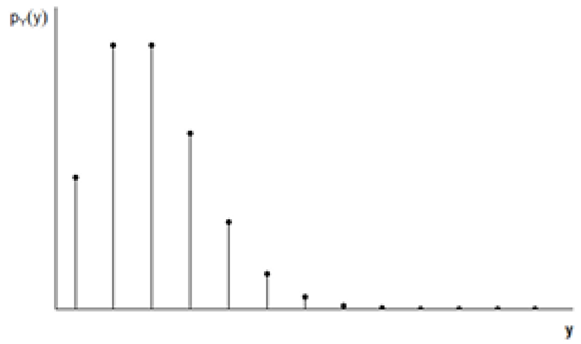


Figura 5.32. *Función de probabilidad puntual Poisson de parámetro* α.

Si $Y \sim Po(\alpha)$, se demuestra que su media es $E(Y) = \alpha$ y su desvío estándar es $D(Y) = \sqrt{\alpha}$.

Dentro de las distribuciones de Poisson existe el proceso Poisson, el cual es un caso particular donde el parámetro α, está multiplicado por t.

Entonces αt será nuestro parámetro (t puede ser el tiempo).

Una familia de variables aleatorias discreta $\{Y_t / t \geq 0\}$ tiene un comportamiento que puede ser descripto por un proceso Poisson con parámetro matemático αt, y se simboliza $Y_t \sim Po(\alpha t)$, si la función de probabilidad para un determinado t es:

$$p_Y(y) = \frac{e^{-\alpha t}(\alpha t)^y}{y!} \quad con\ y = 0, 1, 2, \ldots$$

donde $\alpha \geq 0$.

Si $Y_t \sim Po(\alpha t)$, se demuestra que su media es $E(Y) = \alpha t$ y su desvío estándar es $D(Y) = \sqrt{\alpha t}$.

Si tenemos un proceso Poisson, estamos hablando de una función Discreta, si tomamos el caso del tiempo y este lo planteamos discreto en intervalos de tiempo de 1 hora.

Podemos decir bueno entre los intervalos de una hora, puedo pensar en intervalos de 30 minutos.

Después podés pensar pero en intervalos de 15 minutos

Y así, hasta que sean intervalos infinitesimales, es ahí cuando la distribución se convierte en una exponencial.

Capítulo 6: Muestras aleatorias e inferencia, estimación de parámetros

Estimacion: Si nosotros tenemos una Muestra, y queremos “INFERIR” sobre la población, necesitamos realizar una estimación a partir de los resultados que poseemos.

Existen dos Estimaciones, **PUNTUAL** o **POR INTERVALO DE CONFIANZA**.
Para estimar necesitamos un PARÁMETRO poblacional a estimar con un ESTIMADOR, un resultado que se pueda obtener de la muestra.

Una estimación puntual, consideraremos nuestro parámetro como igual al estimador obtenido.
En una estimación por intervalo de confianza daremos un margen de error, en el cual creemos que está nuestro parámetro, con cierto grado de confianza.

Las pruebas de hipótesis, son resultados que se plantean a partir del parámetro de interés, por ejemplo, que la media poblacional sea igual a tal valor, mayor o menor. Y luego esto lo analizaremos a partir de la estimación si se aprueba o se desecha la hipótesis nula. Y confirmamos la hipótesis alternativa.

Hipótesis nula: Hipótesis que se plantea como cierta o de interés en primera instancia.
Hipótesis alternativa: Hipótesis que contradice y es la contraria a la nula.

Resumen:

Objetivo	¿Con qué herramienta resolverlo?
Estimar un parámetro θ	Intervalo de confianza para θ
Probar una hipótesis sobre algún cambio en θ	Intervalo de confianza para θ Método de 5 pasos Valor p

Los estimadores deben ser:

Insesgados, La esperanza del Estimador debe ser el Parámetro.

Parámetro	Estimador	Esperanza del estimador	
μ	$\overline{X}$	$\mu_{\overline{X}} = E(\overline{X}) = \mu$	El promedio de todos los posibles valores de $\overline{\mathbf{X}}$ es igual al parámetro μ *(Propiedad 1, Material N° 7, página 18)*
σ^2	$\mathbf{S}^2_{\mathbf{n-1}}$	$\mu_{S^2_{n-1}} = E(S^2_{n-1}) = \sigma^2$	El promedio de todos los posibles valores de $\mathbf{S}^2_{\mathbf{n-1}}$ es igual al parámetro $\boldsymbol{\sigma}^2$ *(Propiedad 1, Material N° 7, página 25)*
π	f_r	$\mu_{fr} = E(f_r) = \pi$	El promedio de todos los posibles valores de $\mathbf{f_r}$ es igual al parámetro $\boldsymbol{\pi}$ *(Propiedad 1, Material N° 7, página 22)*

Eficientes, si varios son insesgados se debe elegir el que tenga menor variancia

Consistente, a medida que aumento el tamaño de la muestra, la variancia tiene que tender a 0, lo que nos indicaría que con seguridad estamos teniendo el valor del parámetro a estimar.

Tabla 2.2. Los estimadores $\overline{X}$, f_r y $\mathbf{S}^2_{n-1}$ son consistentes

Parámetro	Estimador	Esperanza del estimador	Desviación estándar del estimador	
μ	$\overline{X}$	$E(\overline{X}) = \mu$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$	La desviación estándar del estimador disminuye cuando n aumenta. Es decir, a medida que el número de observaciones obtenidas aumenta, el promedio de los valores observados se acerca más y más a μ. *(Propiedad 2, Material N° 7, página 17)*
σ^2	$\mathbf{S}^2_{n-1}$	$E(S^2_{n-1}) = \sigma^2$	$\sigma_{S^2_{n-1}} = \sqrt{Var(S^2_{n-1})} = \sqrt{\frac{2\sigma^4}{n-1}}$	La desviación estándar del estimador disminuye cuando n aumenta. Es decir, a medida que el número de observaciones obtenidas aumenta, la variancia de los valores observados se acerca más y más a σ^2 *(Propiedad 2, Material N° 7, página 24)*
π	f_r	$E(f_r)=\pi$	$\sigma_{f_r} = \sqrt{Var(f_r)} = \sqrt{\frac{\pi(1-\pi)}{n}}$	La desviación estándar del estimador disminuye cuando n aumenta. Es decir, a medida que el número de observaciones obtenidas aumenta, la frecuencia relativa muestral se acerca más y más a π. *(Propiedad 2, Material N° 7, página 21)*

Estimación puntual:

Supongamos que queremos hacer una estimación puntual, primero pensamos un Estimador, para el caso de la media poblacional, usamos la media muestral. Luego tomamos la muestra, analizamos su variable Y, y finalmente calculamos el promedio muestral, el cual decimos que la media población debe ser igual a ese por estimación puntual.

Esta es una buena aproximación (estimación) siempre y cuando la muestra que se haya tomado este bien tomado, este valor de media muestral variará aleatoriamente con respecto a las muestras que tomemos.

Estimación por intervalo de confianza:

PASOS A SEGUIR PARA LA CONSTRUCCIÓN DEL INTERVALO DE CONFIANZA

Definir el parámetro a estimar

Seleccionar un buen estimador

Fijar el grado de confianza

Definir una variable "y" que contenga al parámetro a estimar, al estimador y cuya distribución de probabilidad sea conocida (si en la expresión interviene otro parámetro, este debe ser conocido)

Sobre el eje de variación de "y" se eligen dos valores tal que

$$P\left(y_{\alpha/2} < y < y_{1-\alpha/2}\right) = 1-\alpha$$

Despejar el valor del parámetro

Obtener la muestra, a partir de la cual se calcula el intervalo de confianza (a ; b). **Se concluye, con un grado de confianza de (1-α), que el intervalo (a;b) cubre el valor del parámetro.**

Expresiones de algunos intervalos de confianza:

Para el promedio de una población normal con variancia conocida

$$\bar{x} \pm z_{1-\alpha/2} \frac{\sigma_x}{\sqrt{n}}$$

Para el promedio de una población normal con variancia desconocida

$$\bar{x} \pm t_{n-1,1-\alpha/2} \frac{s_x}{\sqrt{n}}$$

Para la variancia de una población normal

$$\left(\frac{(n-1)s_x^2}{\chi^2_{n-1,1-\alpha/2}} ; \frac{(n-1)s_x^2}{\chi^2_{n-1,\alpha/2}} \right)$$

Para una proporción poblacional (para muestras grandes)

$$h \pm z_{1-\alpha/2} \sqrt{\frac{h(1-h)}{n}}$$

En esos 4 casos, tendremos la cota inferior y la superior, con un intervalo de confianza de (1 - α) , el cual buscamos que sea lo más grande posible, ya que queremos que Alpha sea lo más pequeño posible.

Para analizar si una muestra es cumple con una distribución normal, se debe analizar una prueba de linealidad de los percentiles observados vs los percentiles teóricos. Si da una recta, esto indica que la muestra tiene comportamiento Normal.

PRUEBA GRÁFICA DE NORMALIDAD

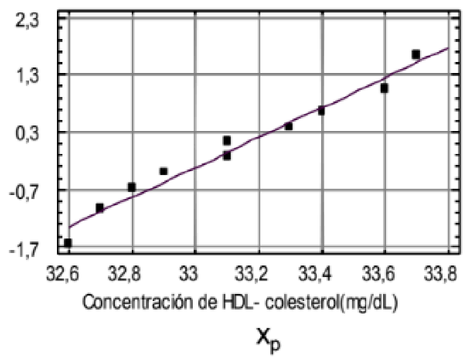
Si $x \longrightarrow N(\mu_x, \sigma_x)$: $$z = \frac{x - \mu_X}{\sigma_X} = -\frac{\mu_X}{\sigma_X} + \frac{1}{\sigma_X} x$$

$\Longrightarrow$ Debe verificarse una relación lineal entre los percentiles z_p y x_p

Percentiles Observados x_p	f_j	F_j	$p = (F_j - 0.5)/n$	Percentiles Teóricos z_p
32.6	1	1	0.05	-1.645
32.7	1	2	0.15	-1.040
32.8	1	3	0.25	-0.670
32.9	1	4	0.35	-0.390
33.1	1	5	0.45	-0.130
33.1	1	6	0.55	0.130
33.3	1	7	0.65	0.390
33.4	1	8	0.75	0.670
33.6	1	9	0.85	1.040
33.7	1	10	0.95	1.645

↓ Valores observados

↓ Frec Rel Acum Corregida



Finalmente, para concluir podemos usar los estimadores para analizar si una hipótesis es falsa o verdadera.

Como un ejemplo, si quisiéramos comparar si una muestra cumple con cierta característica de que el promedio poblacional sea igual a algún valor determinado.

Nuestra hipótesis nula es que, el parámetro sea igual a μ_0 , si la hipótesis alternativa es que en verdad el parámetro sea mayor (en este caso) es decir que μ sea igual a μ_1 siendo $\mu_1 > \mu_0$.

Para esto planteamos nuestro estimador, en este caso el promedio. Y analizamos que valor tiene que tener la cota superior. La cual será, como se ve abajo el c_α .

Si el promedio de nuestra muestra es mayor que nuestra cota superior aceptable, entonces se rechaza la hipótesis nula.

ENSAYO DE HIPÓTESIS PARA EL PROMEDIO DE UNA POBLACIÓN NORMAL CON VARIANCIA CONOCIDA

$H_0)\ \mu = \mu_0$ $\quad$ $H_1)\ \mu = \mu_1$ $\quad$ $\mu_1 > \mu_0$ $\quad$ $P(E_I) = \alpha$

Estadística básica: $\overline{x}$

Se rechaza H_0 cuando:

$$\overline{x} > c_\alpha = \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$

Es decir, pensamos en donde cae nuestra muestra si suponemos que la poblacion tiene una distribución Normal con promedio en μ_0