

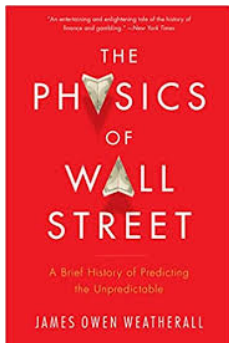
# Estadística 2



Andrés Martínez  
Finanzas

- ▶ Contenido
- ▶ Evaluación
- ▶ Horario de Atención
- ▶ Proyecto Final

■ Puedo predecir con exactitud los movimientos planetarios, pero no puedo predecir la locura de las masas ■ Isaac Newton



**Definición:** Siendo  $\Omega$  un conjunto no vacío con una colección de subconjuntos  $\mathcal{F}$ . Se dice que  $\mathcal{F}$  es una  $\sigma$  álgebra con las siguientes propiedades:

- ▶ El conjunto vacío  $\emptyset$  pertenece a  $\mathcal{F}$
- ▶ Como sea que un conjunto  $\mathcal{A}$  pertenece a  $\mathcal{F}$ ,  $\mathcal{A}^c$  también pertenece a  $\mathcal{F}$ .
- ▶ Como sea que unos conjuntos  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$  pertenecen a  $\mathcal{F}$ , la unión  $\cup$  de  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$  también pertenece a  $\mathcal{F}$ .

**Propiedades:** Siendo  $\Omega$  un conjunto no vacío con una colección de subconjuntos  $\mathcal{F}$ . Una medida de probabilidad  $\mathbb{P}$  es una función que para cada conjunto de  $\mathcal{A} \in \mathcal{F}$  asigna un número entre  $[0, 1]$  llamado probabilidad de  $\mathcal{A}$  y escrito como  $\mathbb{P}(\mathcal{A})$  :

- ▶  $\mathbb{P}(\Omega) = 1$
- ▶ Como sea que  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$  es una secuencia de conjuntos disyuntos en  $\mathcal{F}$
- ▶

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} \mathcal{A}_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(\mathcal{A}_n)$$

Por lo tanto, cualquier variable aleatoria dentro de un proyecto financiero se mide a partir del espacio de probabilidad triple  $(\Omega, \mathcal{F}, \mathbb{P})$ .

**Teorema:** Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Para  $A, B$   $(A_n)_{n \in \mathbb{N}}$  y  $(B_n)_{n \in \mathbb{N}} \in \mathcal{F}$

- ▶  $\mathbb{P}(\emptyset) = 0$
- ▶  $P(A^c) = 1 - P(A)$
- ▶ Si  $A \subseteq B$  entonces  $P(A) \leq P(B)$
- ▶  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

**$\sigma$  algebra:** La dupla  $(\Omega, \mathcal{F})$  se llama espacio medible si  $\Omega \neq \emptyset$  y  $\mathcal{F}$  una  $\sigma$  álgebra, es decir, un sistema de subconjuntos de un conjunto que posee las siguientes propiedades:

- a  $\Omega \in \mathcal{F}$
- b Si  $A \in \mathcal{F}$ , entonces  $\overline{A} := \Omega \setminus A \in \mathcal{F}$ .
- c Si  $A_1, A_2, \dots \in \mathcal{F}$  entonces  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

**Espacio Borel:** Siendo  $X$  una variable aleatoria en un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . La medida de distribución de  $X$  es la medida de probabilidad  $\mu_X$  que asigna a cada subconjunto Borel  $\mathcal{B}$  de  $\mathbb{R}$  la medida  $\mu_X(B) = \mathbb{P}\{X \in B\}$ .

**Definición:** La menor  $\sigma$  álgebra sobre  $\mathbb{R}$  que contiene todos los intervalos de la forma  $(-\infty, a]$  con  $a \in \mathbb{R}$  se llama  $\sigma$  álgebra de Borel y se denota por  $\mathcal{B}$ . Los elementos de  $\mathcal{B}$  se llaman conjuntos de Borel.



## Experimento de Laplace

El experimento de Laplace está compuesto por la probabilidad del conjunto  $A$  que es igual a los elementos de  $A$  dividido en los elementos de  $\Omega$ .

$$P(A) = \frac{|A|}{|\Omega|}$$

**Probabilidad condicionada** Para  $P(A) \neq 0$  existe una probabilidad condicionada de dos subconjuntos en  $\Omega$  por lo tanto la probabilidad condicionada de  $B$  con respecto a  $A$  es igual:

$$P_A(B) \equiv P(B | A) = \frac{P(B \cap A)}{P(A)} \text{ para } P(A) > 0$$

## Independencia de Resultados

Dos conjuntos  $A$  y  $B$  son independientes siempre que:

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

## Probabilidad Total

Si no hay independencia en los resultados se dice que el evento la probabilidad del evento  $B$   $P(B)$  es la suma de todos los eventos que suceden en  $A_1, \dots, A_n$  de esta forma:

$$P(B) = \sum_i^n P(A_i)P(B | A_i)$$

## Teorema de Bayes

Los métodos bayesianos son herramientas de análisis de datos que se derivan de los principios de inferencia bayesiana.

- ▶ Estimaciones de parámetros con buenas propiedades estadísticas
- ▶ Descripciones parsimoniosas de los datos observados
- ▶ Predicciones para datos faltantes y pronósticos de datos futuros
- ▶ Un marco computacional para la estimación, selección y validación de modelos

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_i^n P(A_i)P(B | A_i)}$$

**Definición** Sean  $(\Omega, P)$  un espacio de probabilidad finito. Una variable es una función real definida en  $\Omega$  dos espacios medibles.

- a Una función  $f : \Omega \rightarrow \Omega'$  se llama  $\mathcal{F}$  medible.
- b Si, además, está definida una probabilidad  $P$  sobre  $(\Omega, \mathcal{F})$  entonces  $X$  se mide en un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ .

Una función de densidad  $f(x)$  de una variable aleatoria  $X$  tiene las siguientes propiedades:

a  $f(x) \geq 0$

b  $\int_{-\infty}^{\infty} f(x) dx = 1$

c  $P(a \leq X \leq b) = \int_a^b f(x) dx$

Si se quiere obtener la distribución acumulada de una variable aleatoria  $X$  discreta o continua, se puede usar la función de distribución acumulada que tiene las siguientes propiedades:

- a  $0 \leq F(x) \leq 1$
- b  $F(x_1) \leq F(x_2)$
- c  $\lim_{t \rightarrow -\infty} F(x) = 0$
- d  $\lim_{t \rightarrow \infty} F(x) = 1$
- e  $F(x)$  es continua por la derecha

**Definición** Sean  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad y  $X$  una variable aleatoria numérica.

- a Si  $X$  es una variable aleatoria discreta con valores  $x_1, x_2, \dots$  se dice que la esperanza de  $X$  existe si
- $$\sum_{k=1}^{\infty} |x_k| = P(\{X = x_k\}) < \infty.$$

El valor esperado de una variable aleatoria se obtiene sumando todos sus resultados y multiplicándolo por la probabilidad

$$\mathbb{E}(X) = \sum_x x \mathbb{P}(x) \quad (1)$$

- b Si  $X$  es una variable aleatoria continua con una función de densidad  $f(x)$ , se dice que la esperanza de  $X$  existe si  $\int_{-\infty}^{\infty} |x|f(x)dx < \infty$

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (2)$$

- c Si  $E[X] < \infty$ , entonces  $Var[X] := E([X - E(X)]^2)$  y  $\sigma = \sqrt{Var(X)}$



**Propiedades del Valor Esperado:** Sean  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad y  $X$  una variable aleatoria real.

- ▶  $E(a) = a$  para todo  $a \in \mathbb{R}$ .
- ▶ si  $X_1 \leq X_2$ , entonces  $E(X_1) \leq E(X_2)$
- ▶  $E(aX_1 + bX_2) = aE(X_1) + bE(X_2)$  para todo  $a, b \in \mathbb{R}$
- ▶  $|E(X)| \leq E(|X|)$

**Propiedades de la Varianza:** Sean  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad y  $X$  una variable aleatoria real.

- ▶  $Var(X) = E(X^2) - (E(X))^2$
- ▶  $Var(aX + b) = a^2 Var(X)$  para  $a, b \in \mathbb{R}$

$$STD[X] = \sqrt{V[X]} \equiv \sigma_X$$

**Distribución Binomial** Es una distribución para una variable aleatoria discreta, la probabilidad se obtiene con

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \quad (3)$$

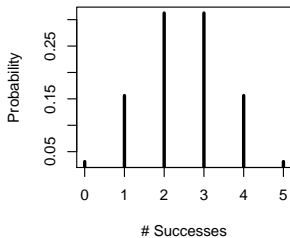
Donde  $p$  y  $q$  son probabilidades  $n$  el número de escenarios y  $x$  el valor que se desea obtener.

El valor esperado y la varianza son:

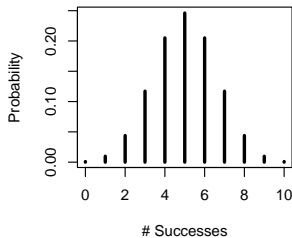
$$\mathbf{E}(X) = np$$

$$\mathbf{V}(X) = np(1 - p)$$

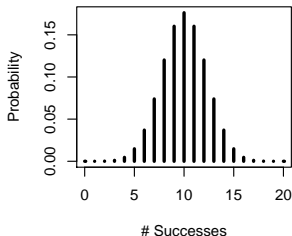
**Binomial Distribution ( $n=5$ ,  $p=0.5$ )**



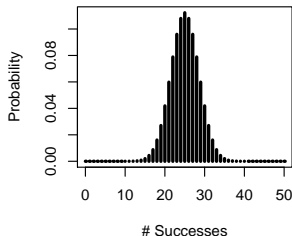
**Binomial Distribution ( $n=10$ ,  $p=0.5$ )**



**Binomial Distribution ( $n=20$ ,  $p=0.5$ )**



**Binomial Distribution ( $n=50$ ,  $p=0.5$ )**



Su función de probabilidad es:

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

Mientras que la función de densidad se define como:

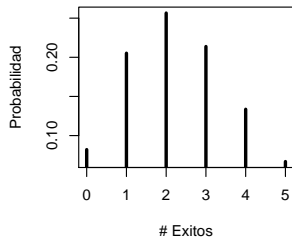
$$F(x) = \sum_{x=0}^n P(x) = e^{-\lambda} \sum_{x=0}^n \frac{\lambda^x}{x!}$$

Con un valor esperado y una varianza de:

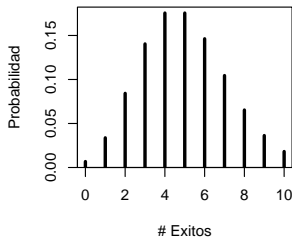
$$\mathbf{E}(X) = \lambda$$

$$\mathbf{V}(X) = \lambda$$

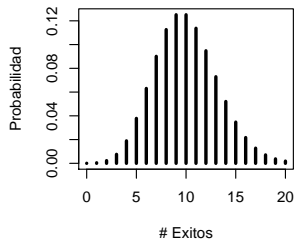
**Distribución Poisson lambda = 2.5**



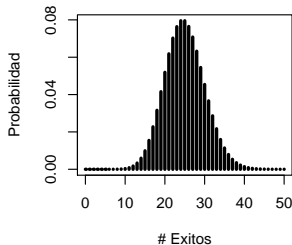
**Distribución Poisson lambda = 5**



**Distribución Poisson lambda = 10**



**Distribución Poisson lambda = 25**



La distribución mas sencilla de la teoría de probabilidad, se encuentra entre dos intervalos  $[a, b]$  y su distrubición igual se denota de la forma  $X \sim \mathbb{U}(a, b)$  lo que quiere decir que esta función se distribuye de forma uniforme con parámetros  $a$  y  $b$  que son constantes en el espacio cartesiano.

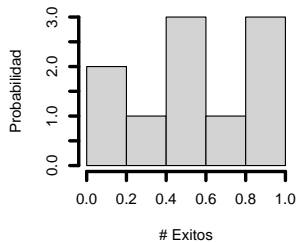
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } a \leq x \leq b \\ 0 & \text{de otra forma} \end{cases}$$

El valor esperado y la varianza son:

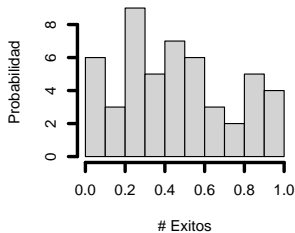
$$\mathbf{E}(X) = \frac{b + a}{2}$$

$$\mathbf{V}(X) = \frac{(b - a)^2}{12}$$

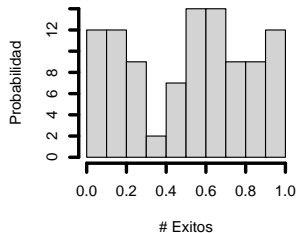
**Distribución Uniforme n=10**



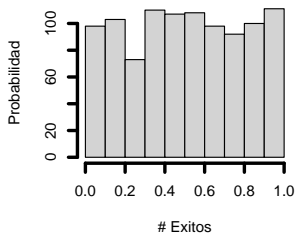
**Distribución Uniforme n=50**



**Distribución Uniforme n=100**



**Distribución Uniforme n=1000**



Es la más importante de las distribuciones continuas. Su característica principal es la simetría, ya que la media, la mediana y la moda tienen el mismo valor; su sesgo es cero y su curtosis es 3, lo que significa que es poco probable que presente escenarios extremos.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (\mu \in (-\infty, \infty), \sigma > 0)$$

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

Función de distribución acumulada

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) du$$



$$\mu = 0, \sigma = 1: X \sim N(0; 1).$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x)^2}{2}\right) (\mu \in (-\infty, \infty), \sigma > 0)$$

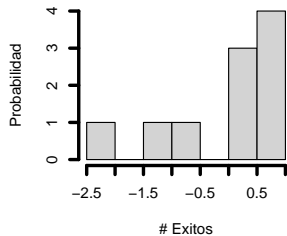
$$X \sim Z(0; 1)$$

Función de distribución acumulada

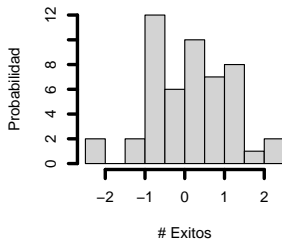
$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(\frac{-(x)^2}{2}\right) du$$

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

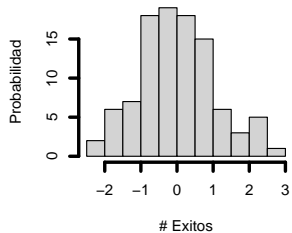
**Distribución Normal  $n=10$**



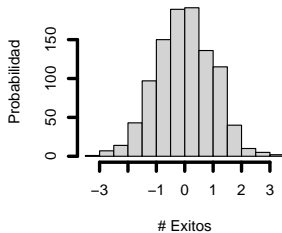
**Distribución Normal  $n=50$**



**Distribución Normal  $n=100$**



**Distribución Normal  $n=1000$**



Los precios de los instrumentos financieros tendrán este tipo de distribución, y esta es una de las razones por las cuales se trabaja mejor con rendimientos dado que este tipo de medición es más complejo. El resultado de la variable siempre es positivo como en los precios de los activos  $x > 0$  y la variable aleatoria se distribuye  $X \sim LN(0; 1)$ .

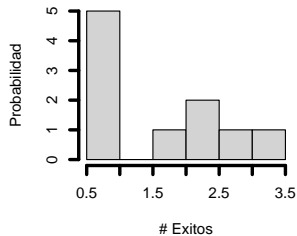
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{x} \exp\left(-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right)$$

Para toda variable aleatoria que se distribuye de forma normal,  $X \sim LN(\mu; \sigma^2) \leftrightarrow \exp(Y) = X \sim LN(\mu; \sigma^2)$ .

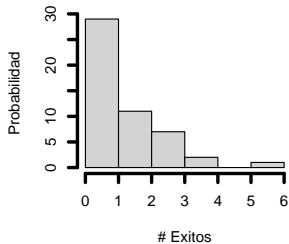
$$\mathbf{E}(X) = e^{\mu + \frac{\sigma^2}{2}},$$

$$\mathbf{V}(X) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1),$$

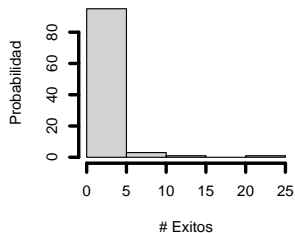
**Distribución Log Normal n=10**



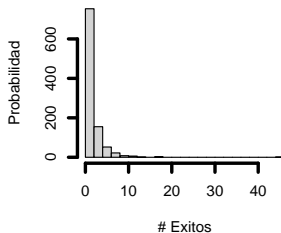
**Distribución Log Normal n=50**



**Distribución Log Normal n=100**



**Distribución Log Normal n=1000**



Este curso parte del supuesto de normalidad para medir y analizar el riesgo financiero, desde la estadística descriptiva se puede evaluar de la siguiente forma

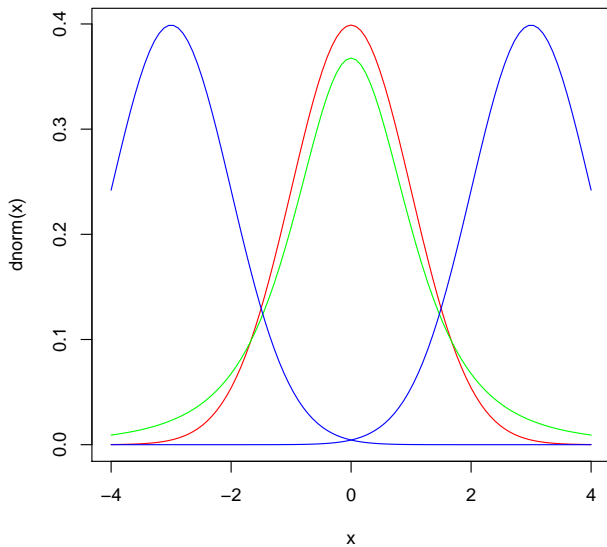
$$Sesgo = \frac{\sum_{i=1}^n (x_i - \mu)^3}{(n-1)\sigma^3} \quad (4)$$

$$Kurtosis = \frac{\sum_{i=1}^n (x_i - \mu)^4}{(n-1)\sigma^4} \quad (5)$$

El estadístico se obtiene con ayuda de la ecuación 4 y 5

$$LM = N \left[ \frac{Sesgo^2}{6} + \frac{(Kurtosis - 3)^2}{24} \right] \quad (6)$$

## Forma de la distribución





## Métodos de muestreo

La distribución de probabilidad asociada con la población (de la cual muestra) se llama distribución de población y distribución de probabilidad asociado con nuestra estadística se llama distribución muestral; claramente, los dos están interrelacionados. Para conocer la distribución de la población, es imperativo saber todo lo que podamos sobre la distribución muestral. Por lo tanto, nos enfocamos en lo que sucede en el caso especial del muestreo del distribución normal, y en particular, cumplimos con la distribución muestral de  $\bar{x}$  y  $s^2$ .



Usualmente cuando se lleva a cabo una investigación de forma cuantitativa, es necesario contar con variables que permitan generar las pruebas de contraste de hipótesis. Estas variables provienen de observaciones que se obtienen de una muestra dado que es muy difícil trabajar con una población.

- ▶ Población - Parámetros
- ▶ Muestra - Estadístico
- ▶ error de muestra: Diferencia entre el estadístico y el parámetro

Muestreo de acuerdo a:

- ▶ Objeto de estudio
- ▶ Método utilizado (Paramétricos - No paramétricos)
- ▶ Información considerada (Clásico - Bayesiano)

## Procedimientos de Inferencia

Objetivo	¿Cómo se distribuye $y$ ?	¿Influye $x$ en $y$ ?
Datos	Muestreo $(y_1, \dots, y_n)$	Diseño $x_1 : (y_{1,1}, \dots, y_{1,n})$
Método	Paramétrico	No paramétrico
Información	Clásico - Bayesiano	
Resultados	$\hat{\theta} \in (a, b) - \theta \hat{\mid} y$	Distribución empírica

- ▶ Muestreo aleatorio simple:
  - ▶ Cada elemento de la población tiene la misma probabilidad de ser elegido.
  - ▶ Las observaciones se realizan con reemplazamiento, de esta manera la población es idéntica en todas las observaciones.
- ▶ Muestreo estratificado: Los elementos de la población se dividen en clases o estratos, pero de forma interna la información es homogénea.
- ▶ Muestreo por Conglomerados: En los conglomerados no hay diferencias entre los grupos, pero la información interna de cada grupo es heterogénea.
- ▶ Muestreo sistemático: Variación del aleatorio simple.  $N/n$  con  $k$  como su entero más próximo.

**Teorema 1** Para cualquier población, el valor promedio de todas las posibles medias muestrales calculado a partir de todas las posibles muestras aleatorias de un tamaño dado de la población será igual a la media de la población,  $\mu_{\bar{x}} = \mu$

Estimador insesgado: característica de ciertas estadísticas en las que promedio de todos los valores posibles de la estadística de muestra es igual al parámetro

**Teorema 2** Para cualquier población, la desviación estándar de las posibles medias muestrales calculado a partir de todas las muestras aleatorias posibles de tamaño  $n$  es igual al desviación estándar de la población dividida por la raíz cuadrada del tamaño de la muestra. También llamado error estándar,

$$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}$$

**Teorema 3** Si una población tiene una distribución normal, con media  $\mu$  y un estándar desviación  $\sigma$ , la distribución muestral de la media muestral  $\bar{x}$  también es normalmente distribuida con una media igual a la media de la población ( $\mu_{\bar{x}} = \mu$ ) y una desviación estándar igual a la desviación estándar de la población dividida por la raíz cuadrada del tamaño de la muestra  $\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}$

### Ley de los grandes números

La ley de los grandes números es una propiedad que explica como el incremento del número de experimentos permite que el valor esperado de las simulaciones, se acerque al valor esperado real.

**Definición:** Sea  $X_n$  para  $n \in N$  una sucesión de variables aleatorias reales e integrables sobre un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . Se dice que esta sucesión satisface la ley débil de los grandes números si

$$\frac{1}{n} \sum_{j=1}^n [X_j - E(X_j)] \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \quad (7)$$

**Teorema del Límite Central:** Si  $X_1, X_2, \dots$ , son variables aleatorias reales, independientes e idénticamente distribuidas iid con  $E(X_n^2) < \infty$ . Sean  $\mu := E(X_n)$  y  $\sigma^2 := \text{Var}(X_n)$ . Si  $\sigma > 0$  entonces

$$Z_n := \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} N(0, 1) \quad (8)$$

Sin embargo, hay muchas poblaciones que no son normales. Estadístico a menudo se encuentran tomando muestras de esas poblaciones. ¿Qué se puede decir en este caso? La sorprendente respuesta está contenida en el Teorema del límite central

- ▶ El tamaño de la muestra debe ser "suficientemente grande"
- ▶ Si la población es bastante simétrica, entonces tamaños de muestra tan pequeños como 2 o 3 puede proporcionar una distribución de muestreo normalmente distribuida
- ▶ Si la población está muy sesgada o tiene una forma irregular, la el tamaño de muestra requerido será mayor
- ▶ Una definición conservadora de un tamaño de muestra suficientemente grande es  $n \geq 30$



La distribución muestral de la media se puede obtener de dos formas. La primera cuando conocemos la varianza  $\sigma^2$  en donde podemos decir que:

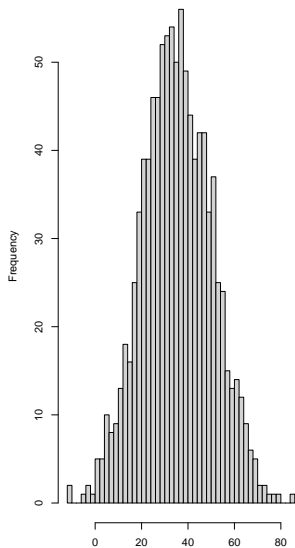
$$\overline{X} \xrightarrow{d} N\left(\mu, \frac{\sigma^2}{n}\right) = \frac{\overline{X}_n - \mu}{\sigma_n / \sqrt{n}} \quad (9)$$

Observe que la varianza depende del tamaño de la muestra  $n$ . En caso de que se desconosca la varianza  $\sigma^2$ , la distribución  $t$  de student ofrece una solución para medir la probabilidad de la media muestral  $\overline{X}$

$$\frac{\overline{X}_n - \mu}{S_n / \sqrt{n}} \xrightarrow{d} t(n - 1) \quad (10)$$

```
> # Ejemplo para de muestreo aleatorio simple
>
>
> n <- 20                                     #
> nreps <- 10000                             #
> sample.mean <- numeric(nreps)             #
> sample.normal <- rnorm(n = 1000, mean = 35, sd = 15)
> for (i in 1:nreps) {
+   sample <- rnorm(n = n, mean = 35, sd = 15)
+
+   sample.mean[i] <- mean(sample)
+
+ }
> hist(sample.normal, breaks = 50, main = "Distribution of
> hist(sample.mean, breaks = 50, main = "Distribution of Sa
```

Distribución Poblacional



Distribución media muestral

