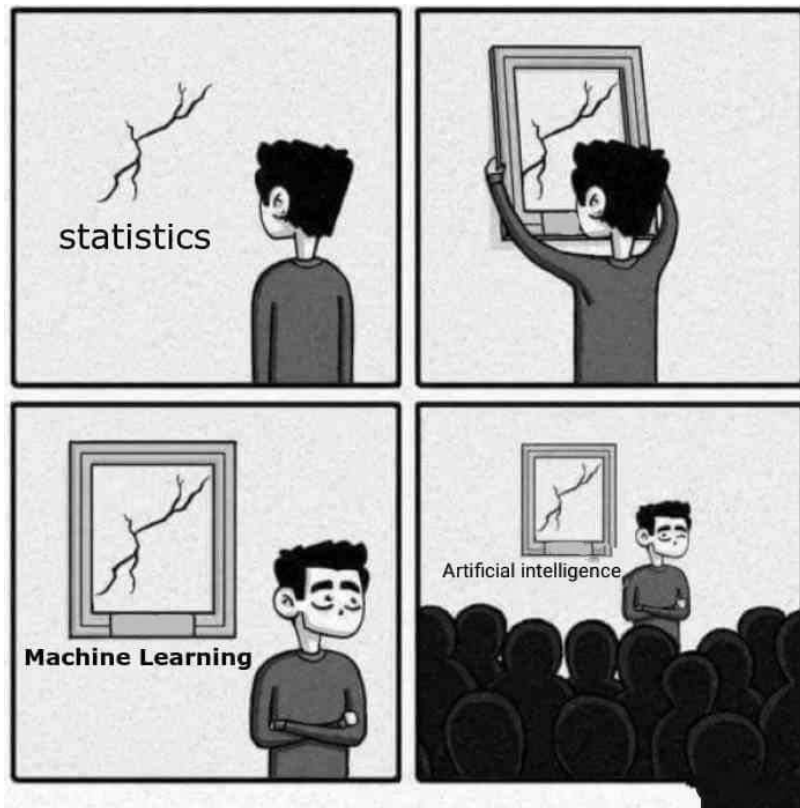
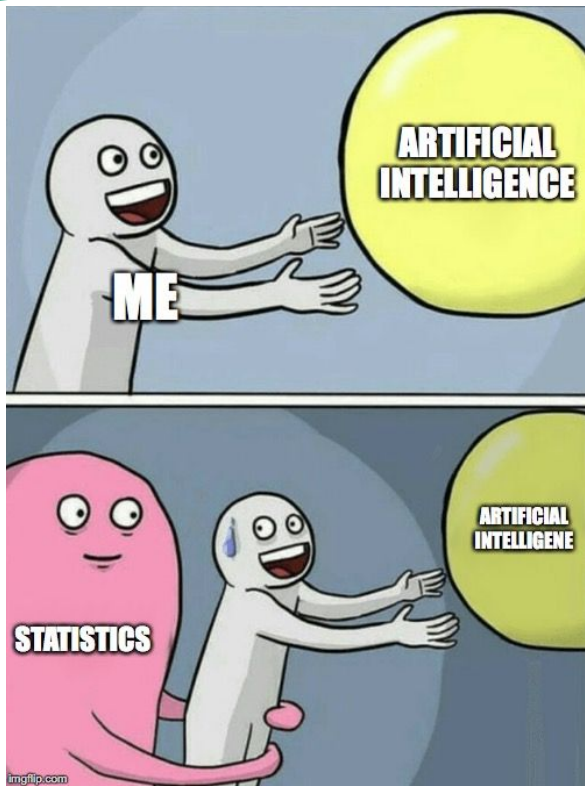


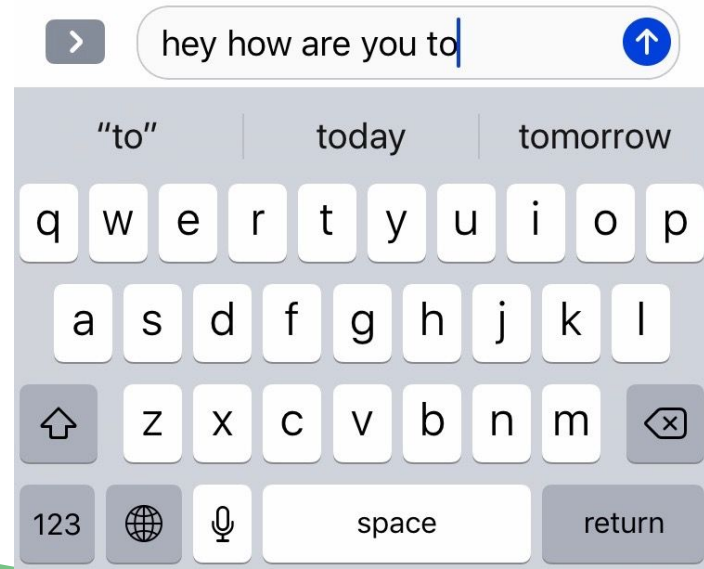
Probabilidad y Estadística

Clase 1

Ud. se encuentra
aquí:

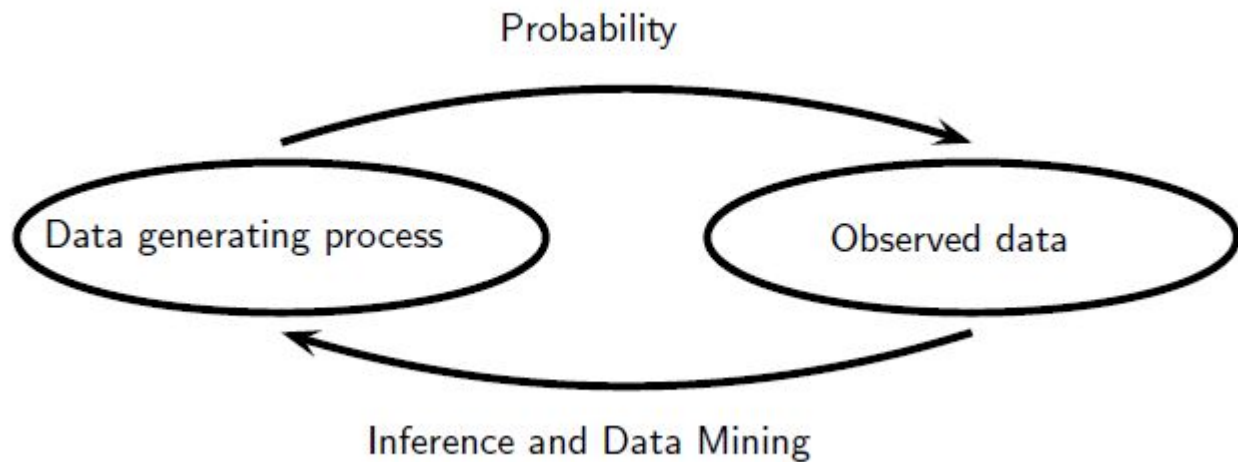
para poder estar
aquí:





Ejemplos de probabilidad en IA?

Probabilidad vs. Estadística

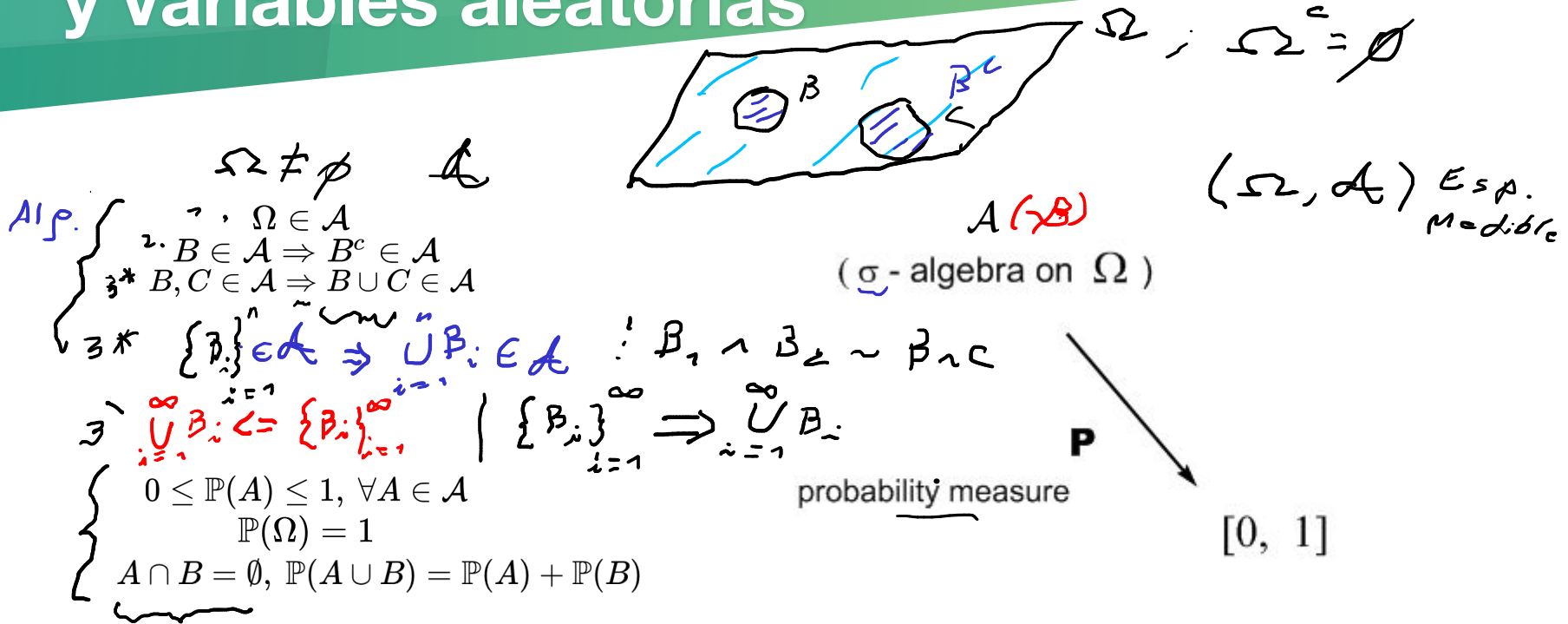


Cronograma

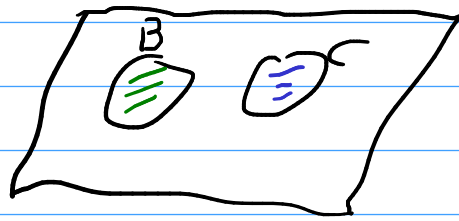
✓	Clase 1	Repaso Distribuciones útiles.
	Clase 2	V.a. condicionadas Esperanza condicional
	Clase 3	ECM y estimadores de cuadrados mínimos Estimador de máxima verosimilitud
✓	Clase 4	Estimación Bayesiana
	Clase 5	Estimación no paramétrica
	Clase 6	Intervalos de confianza
✓	Clase 7	Test de hipótesis
	Clase 8	Examen

Repaso

Espacio de Probabilidad (Ω, \mathcal{A}, P) y variables aleatorias

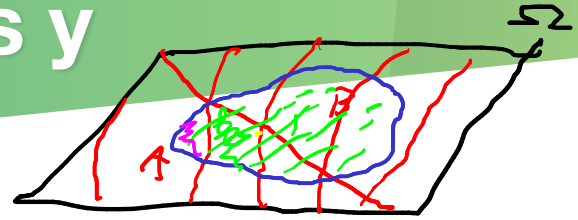


$$3) B, C \in \mathcal{L} \Rightarrow B \cup C \in \mathcal{L}$$



Probabilidades condicionales y proba. total

$$\mathbb{P}(A \cap B_1) \cup \dots \cup (A \cap B_n)$$



Def: Se llama **probabilidad condicional** de A dado B ($\mathbb{P}(A|B)$) a la probabilidad de que ocurra el evento A sabiendo que B ha ocurrido, y está definida por

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

Def: Diremos que los eventos B_1, \dots, B_n forman una **partición** si $B_i \cap B_j = \emptyset \forall i, j$ $i \neq j$ y $\bigcup_{i=1}^n B_j = \Omega$.

Luego podemos describir al evento A como $A = (A \cap B_1) \cup \dots \cup (A \cap B_n)$ de forma que

$$\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \dots + \mathbb{P}(A \cap B_n) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \dots + \mathbb{P}(A|B_n)\mathbb{P}(B_n)$$

Fórmula de probabilidad total

Teorema de Bayes e independencia

Teorema de Bayes: Sean B_1, \dots, B_n una partición de Ω , y A un evento con probabilidad positiva:

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{i=1}^n \mathbb{P}(A|B_i)\mathbb{P}(B_i)} = \mathbb{P}(A) > 0$$

"Test 99%"
Test ~ 0,19

Def: Diremos que dos eventos A y B son **independientes** si y sólo si vale que

$$A, B \text{ son Ind.} \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

$$A, B \text{ son Ind.} \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

$$(\Omega, \mathcal{A})$$

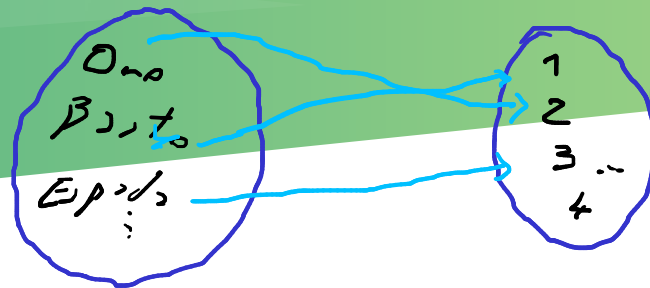
G-alp.

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

$\omega \mapsto X(\omega)$

Variables aleatorias

Variables aleatorias



Una va. X es una función que mapea eventos a números.

Ejemplo: basto $\rightarrow 1$, Oro $\rightarrow 2$, espada $\rightarrow 3$, copa $\rightarrow 4$

X tiene asociada una función de distribución, definida como

$$X \sim F_X$$

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

además se define la función de distribución inversa o función cuantil como

$$F_X^{-1}(q) = \inf\{x : F(x) > q\}$$

$$F_X(x) \in [0, 1] \forall x \in \mathbb{R}$$

$F_X(x)$ es monótona no decreciente

$F_X(x)$ es continua por derecha

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad \text{y} \quad \lim_{x \rightarrow \infty} F_X(x) = 1$$

„

Si F_X es estrictamente creciente y continua, F_X^{-1} es su inversa.

Tipos de v.a.

- ① • Discretas (v.a.d): toman valores en un conjunto discreto o numerable de puntos. Si X es **v.a.d**, tendrá además función de probabilidad dada por

$$p_X(x) = \mathbb{P}(X = x)$$

- ② • Continuas (v.a.c): toman valores en un intervalo continuo. Si X es una **v.a.c**, tendrá asociada una función de densidad

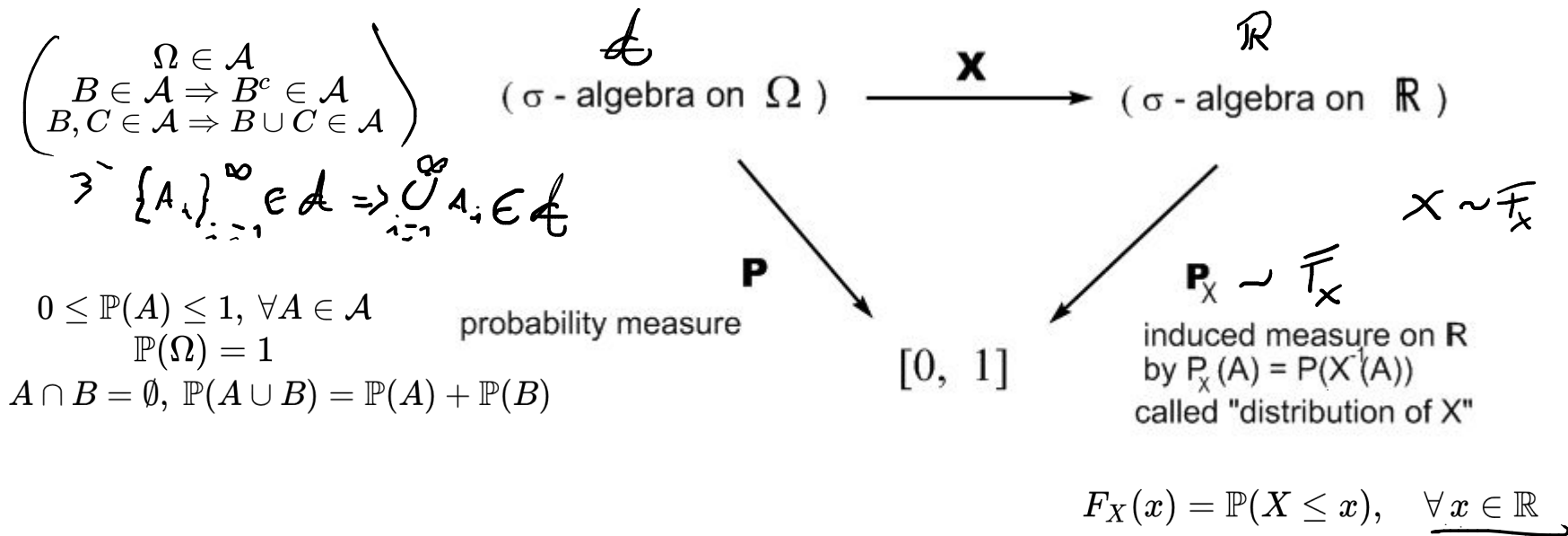
$$f_X(x) = \frac{dF_X(x)}{dx} \quad F_X(x) = \int_{-\infty}^x f_X(u) du$$

¿Qué propiedades debían cumplir cada una de estas funciones?

$$f_x(x) \geq 0$$

$$; \int_{-\infty}^{\infty} f_x(u) du = \underline{1}$$

Espacio de Probabilidad (Ω, \mathcal{A}, P) y variables aleatorias



$$X \sim \mathcal{F}_x$$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$X: \Omega \longrightarrow \mathbb{R}^2$$

$$\omega \longmapsto X(\omega)$$

Vectores aleatorios

Distribución conjunta y marginales

	Y = 0	Y = 1	
X=0	1/10	2/10	3/10
X=1	3/10	4/10	7/10
	4/10	6/10	1

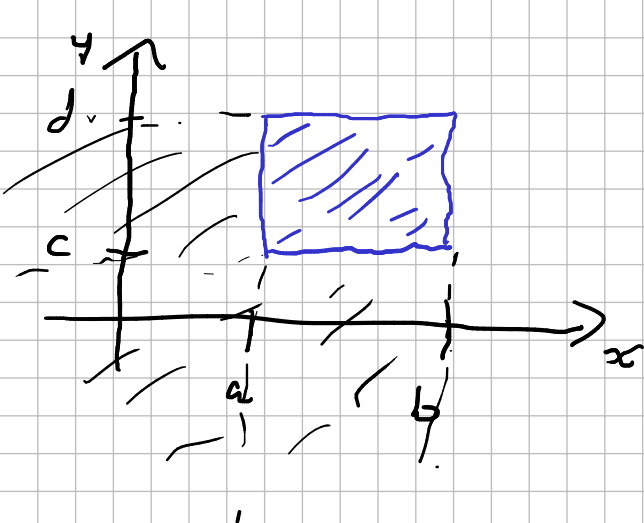
Handwritten notes: $p_{x,y}$ next to the joint probabilities, p_x next to the row sums, and p_y next to the column sums.

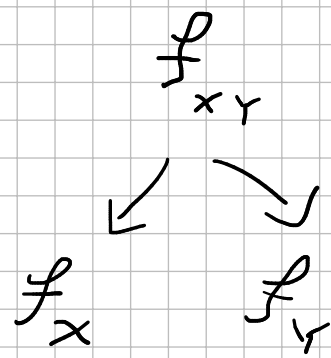
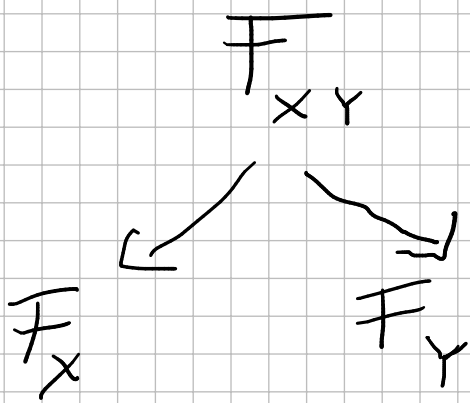
Si tenemos dos variables X e Y se define su **función de distribución conjunta** como $F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y)$ y vale la regla del rectángulo:

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F_{X,Y}(b, d) - F_{X,Y}(a, d) - F_{X,Y}(b, c) + F_{X,Y}(a, c)$$

Caso continuo: $f_{X,Y}(x, y)$ es la **función de densidad conjunta** y se definen las **funciones de densidad marginales** como $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y) dy$ y $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y) dx$

Caso discreto: $p_{X,Y}(x, y)$ es **función de probabilidad conjunta** y se definen las **funciones de probabilidad marginal** como $p_X(x) = \sum_y p_{X,Y}(x, y)$ y $p_Y(y) = \sum_x p_{X,Y}(x, y)$





Independencia de v.a.

Diremos que dos v.a. X e Y son independientes si vale que

$$F_{X.Y}(x, y) = F_X(x)F_Y(y) \quad \forall x, y \in \mathbb{R}$$

Caso **discreto**:

$$p_{X.Y}(x, y) = p_X(x)p_Y(y) \quad \forall x, y \in \mathbb{R}$$

Caso **continuo**:

$$f_{X.Y}(x, y) = f_X(x)f_Y(y) \quad \forall x, y \in \mathbb{R}$$

*En el caso de que la intersección sea 0,
La independencia implicaba intersección cero pero no al revés.*

Momentos: La varianza, La Esperanza Matemática

Momentos

Momentos

$$E(X) = \int x f_X(x) dx$$

Esperanza (o media):

$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$

Caso Disc. = $\sum g(x) p_X(x)$

Varianza:

$$var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}(X)^2$$

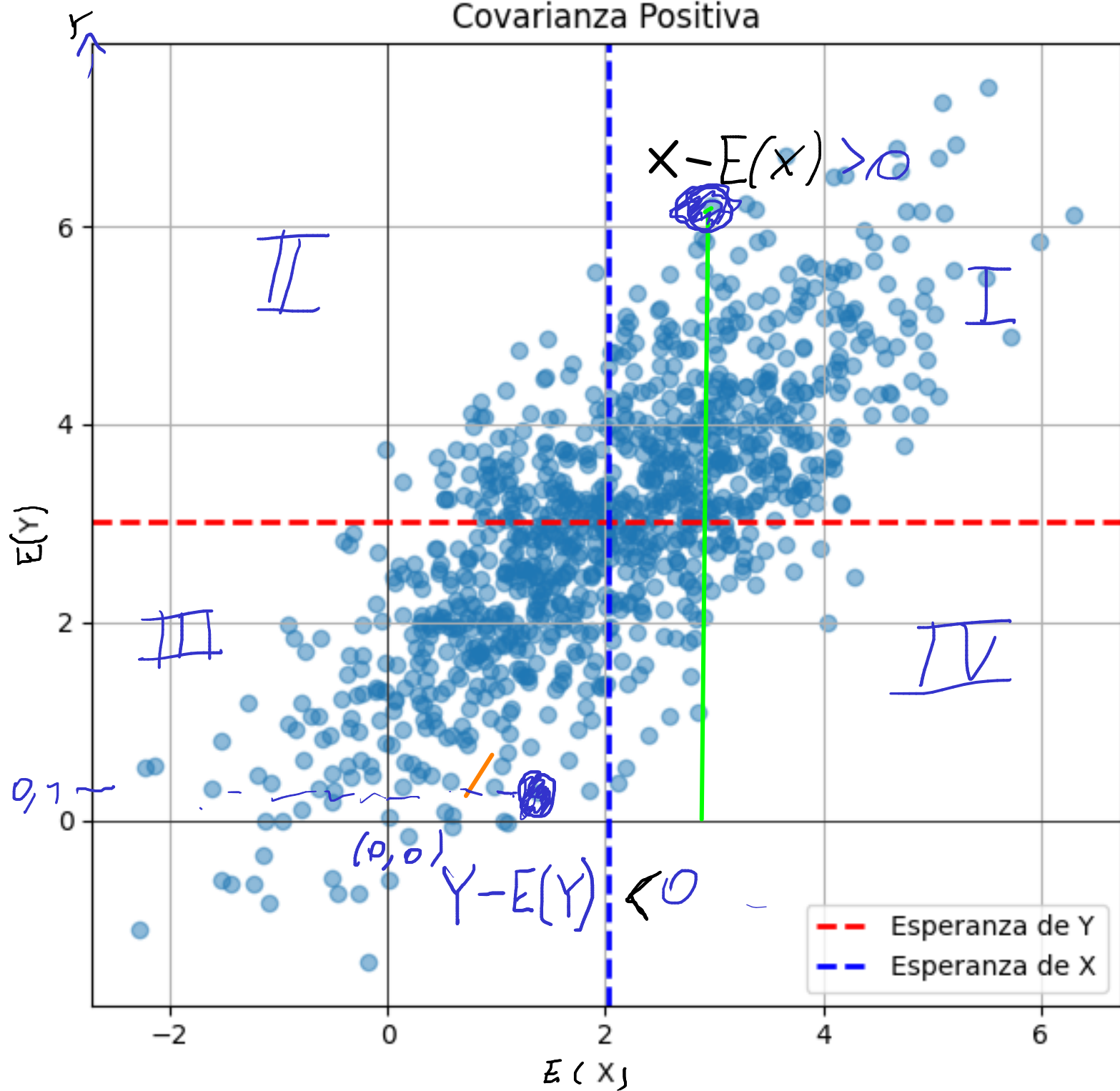
Covarianza:

X, Y

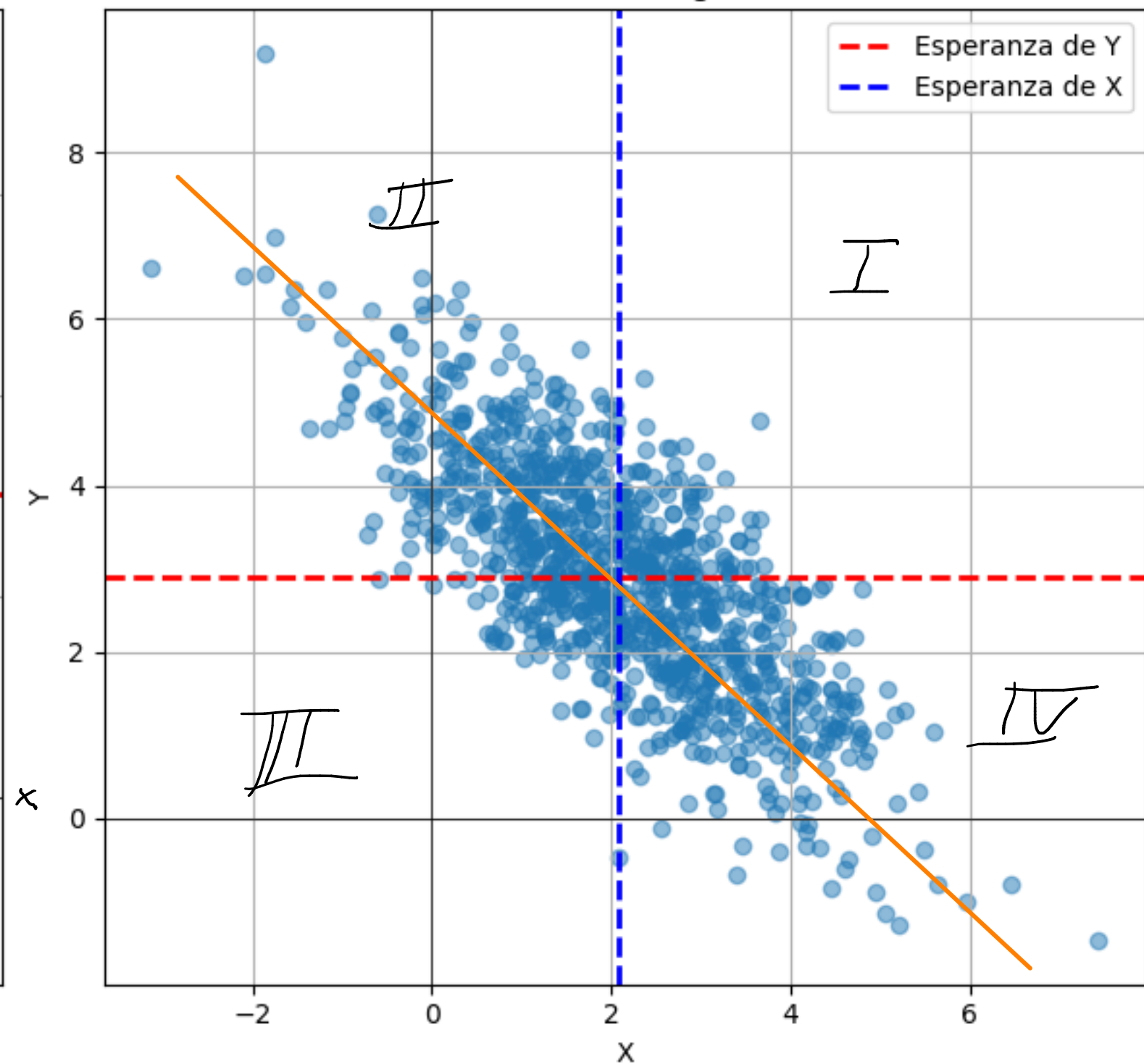
$cov(X, Y) > 0$
 $cov(X, Y) < 0$

$$cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}(X)\mathbb{E}(Y)$$

Covarianza Positiva



Covarianza Negativa



Algunas distribuciones útiles

$$X \sim F_x$$

Variables discretas

$$P(X=x) = p^x \underbrace{(1-p)^{1-x}}_1$$

$$P(X=1) = p \wedge P(X=0) = (1-p)$$

$X \sim$ • **Bernoulli(p)**: $X = \{0,1\}$. Asociada a la ocurrencia o no de un éxito

$Y \sim$ • **Binomial(n, p)**: cantidad de éxitos en n ensayos. $n \geq x$
$$\binom{n}{x} p^x (1-p)^{n-x}$$

$Z \sim$ • **Geométrica(p)**: cantidad de ensayos que debo realizar hasta observar el primer éxito

$$P(1-p)^{x-1}$$

Ejercicio 0



$$p = \frac{1}{6}$$

$$P(X=1) + P(X=2)$$

Se tira sucesivamente un dado

1. Probabilidad de necesitar menos de 3 tiros hasta ver el primer 2
2. Probabilidad de ver exactamente un 2 en 5 tiros

$$1- X \sim G\left(\frac{1}{6}\right)$$

$Z \sim$ • **Geométrica(p)**: cantidad de ensayos que debo realizar hasta observar el primer éxito

$$1/6 + 5/6 * 1/6$$

$$P(1-p)^{x-1}$$

$$2- Y \sim \text{Binomial}(5, \frac{1}{6}) \quad ; \quad p$$

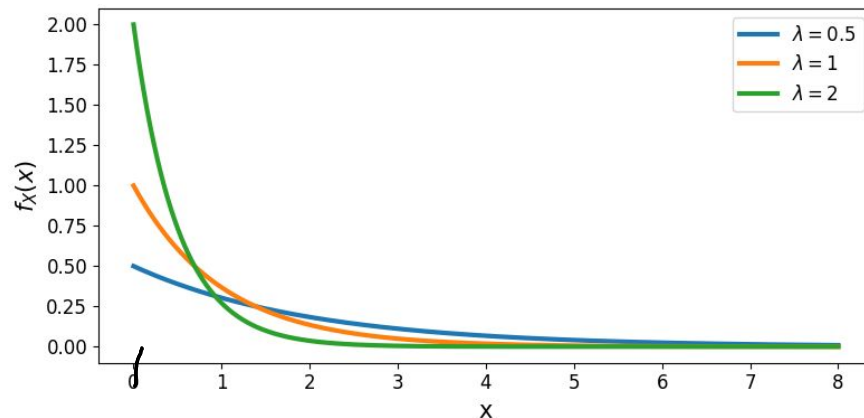
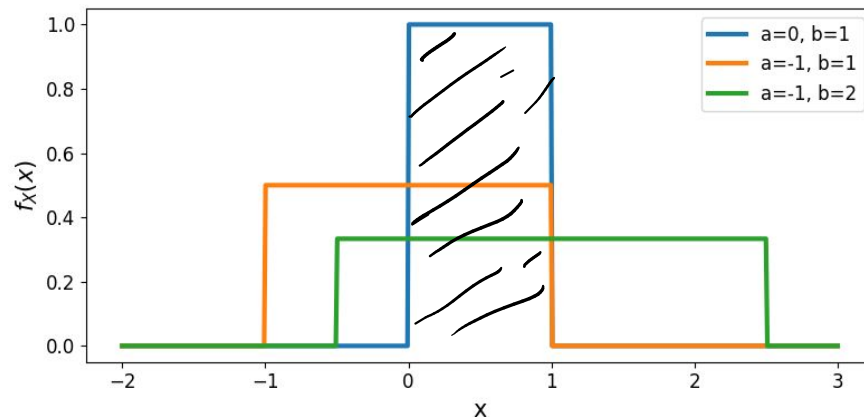
Variables continuas

- **Uniforme:** todos los puntos son equiprobables. $X \sim \mathcal{U}(a, b)$

$$f_X(x) = \frac{1}{b-a} \mathbf{I}\{a < x < b\}$$

- **Exponencial:** sirve para modelar tiempos hasta eventos que no tienen memoria. $X \sim \mathcal{E}(\lambda)$

$$f_X(x) = \lambda e^{-\lambda x} \mathbf{I}\{x > 0\}$$



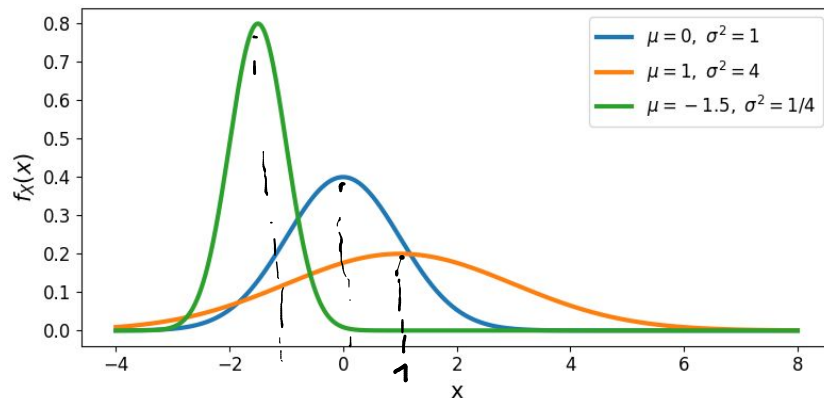
Variables continuas

- Normal (gaussiana).

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

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

μ es la media
 σ^2 es la
varianza



Propiedades: Sean X, Y dos v.a. Independientes

$$X \sim \mathcal{N}(\mu, \sigma^2) \rightarrow \boxed{\frac{X-\mu}{\sigma}} \sim \mathcal{N}(0, 1) \quad (\text{estandarización})$$

$$\underbrace{X \sim \mathcal{N}(\mu_X, \sigma_X^2), Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)}_{\text{(combinación lineal de normales es normal)}} \rightarrow \underbrace{aX + bY}_{\mu} \sim \mathcal{N}(\underbrace{a\mu_X + b\mu_Y}_{\mu}, \underbrace{a^2\sigma_X^2 + b^2\sigma_Y^2}_{\sigma^2})$$

Ejercicio 1

Sea X una v.a con distribución normal estándar. Hallar la probabilidad de que

1. $P(X > 1) = 1 - P(X \leq 1)$
2. $X < -1$
3. $|X| < 2$
4. Hallar los cuantiles 0.1 y 0.9

Sea además $Y \sim N(2, 9)$

1. Hallar $P(2X + Y < 5)$

Ejercicio 2

El tiempo (en minutos) entre llamadas a un call center tiene una distribución exponencial de parámetro $1/5$.

1. Calcular la probabilidad de que la primer llamada llegue después de 2 minutos
2. Calcular la probabilidad de que la primer llamada llegue después de los 5 minutos, si se sabe que en los primeros 3 minutos no se recibieron llamados

Distribución normal multivariada

Función de densidad conjunta

Sea $\underline{X} = [X_1, \dots, X_n]^T$ un vector aleatorio continuo, diremos que X tiene distribución normal multivariada, si su función de densidad conjunta es de la forma

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})}$$

media

matriz de covarianza .

$$\underline{\mu} = [\mu_1, \dots, \mu_n]^T \quad \Sigma = \begin{bmatrix} \sigma_1^2 & cov(X_1, X_2) & \dots & cov(X_1, X_n) \\ cov(X_2, X_1) & \sigma_2^2 & \dots & cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, X_1) & cov(X_n, X_2) & \dots & \sigma_n^2 \end{bmatrix}$$

Notación: $\underline{X} \sim \mathcal{N}(\underline{\mu}, \Sigma)$

Distribuciones marginales

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma) \Rightarrow X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$$

$$\boldsymbol{\mu} = [\mu_1, \dots, \mu_n]^T$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & cov(X_1, X_2) & \dots & cov(X_1, X_n) \\ cov(X_2, X_1) & \sigma_2^2 & \dots & cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, X_1) & cov(X_n, X_2) & \dots & \sigma_n^2 \end{bmatrix}$$

Ejercicio 3

Sean X, Y dos v.a. Con función de densidad conjunta

$$f_{X,Y}(x, y) = \frac{1}{2\pi 0.6} e^{-\frac{1}{2} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix}}$$

1. Calcular $E[X]$, $E[Y]$, $\text{var}(X)$, $\text{var}(Y)$, y $\text{cov}(X, Y)$
2. Hallar las densidades marginales de X e Y
3. Calcular $P(X < 2, Y < -1)$

Bibliografía

Bibliografía

“Mathematical Statistics with Applications”, Dennis D. Wackerly, William Mendenhall III, Richard L. Scheaffer.

“All of Statistics: A Concise Course in Statistical Inference”, Larry Wasserman.