Análisis de Datos para Inteligencia Artificial

Ing. Magdalena Bouza, Dr. Ing. Pablo Briff, Lic. Nicolás Horro

Laboratorio de Sistemas Embebidos - FIUBA

bouza.magdalena@gmail.com, pbriff@fi.uba.ar, nhorro@gmail.com

Mayo de 2021

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Tabla de Contenidos I

Tabla de Contenidos II

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Tabla de Contenidos III

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Tabla de Contenidos V

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Acerca de los docentes

- Magdalena Bouza

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff
- Ingeniero electrónico, FIUBA, 2006

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff
- Ingeniero electrónico, FIUBA, 2006
- Doctor en ingeniería electrónica, FIUBA, 2015

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff
- Ingeniero electrónico, FIUBA, 2006
- Doctor en ingeniería electrónica, FIUBA, 2015
- Email: pbriff@fi.uba.ar

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff
- Ingeniero electrónico, FIUBA, 2006
- Doctor en ingeniería electrónica, FIUBA, 2015
- Email: pbriff@fi.uba.ar

- Nicolás Horro

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff
- Ingeniero electrónico, FIUBA, 2006
- Doctor en ingeniería electrónica, FIUBA, 2015
- Email: pbriff@fi.uba.ar

- Nicolás Horro
- Licenciado en Sistema de Información, Universidad FASTA, 2012

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Acerca de los docentes

- Magdalena Bouza
- Ingeniero electrónico, FIUBA, 2015
- Doctor en ingeniería electrónica, FIUBA, estimado Nov. 2021
- Email: bouza.magdalena@gmail.com

- Pablo Briff
- Ingeniero electrónico, FIUBA, 2006
- Doctor en ingeniería electrónica, FIUBA, 2015
- Email: pbriff@fi.uba.ar

- Nicolás Horro
- Licenciado en Sistema de Información, Universidad FASTA, 2012
- Email: nhorro@gmail.com

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Clase 1: Introducción a Git, Python, Numpy, SciPy y Pandas. Visualización de datos para Machine Learning.

Programa de la Materia

- Clase 1: Introducción a Git, Python, Numpy, SciPy y Pandas. Visualización de datos para Machine Learning.
- Clase 2: Análisis básico de media, desvío estándar, oblicuidad (skewness), curtosis, cuantiles, IQR.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Programa de la Materia

- Clase 1: Introducción a Git, Python, Numpy, SciPy y Pandas. Visualización de datos para Machine Learning.
- Clase 2: Análisis básico de media, desvío estándar, oblicuidad (skewness), curtosis, cuantiles, IQR.
- Clase 3: Variables Aleatorias. Teoría de la información. Entropía. Entropía cruzada. Divergencia KL. Ejemplos. Tests estadísticos univariados para Machine Learning. Introducción a test estadísticos, definición de p-value, z score. Tests de normalidad. Tests de correlación. Tests de independencia. Análisis de varianza (ANOVA). Ejemplos.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Programa de la Materia

- Clase 1: Introducción a Git, Python, Numpy, SciPy y Pandas. Visualización de datos para Machine Learning.
- Clase 2: Análisis básico de media, desvío estándar, oblicuidad (skewness), curtosis, cuantiles, IQR.
- Clase 3: Variables Aleatorias. Teoría de la información. Entropía. Entropía cruzada. Divergencia KL. Ejemplos. Tests estadísticos univariados para Machine Learning. Introducción a test estadísticos, definición de p-value, z score. Tests de normalidad. Tests de correlación. Tests de independencia. Análisis de varianza (ANOVA). Ejemplos.
- Clase 4: Datos, características (features) e Ingenieria de features: Tipos de variables de entrada y salida: continua y categórica (nominal y ordinal).

Programa de la Materia

- Clase 5: Codificación one-hot y dummy. Otros tipos de codificación (binaria, hashing, etc). Valores faltantes (missing values). Normalización de datos. Transformación de datos (box-cox transformation).

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Programa de la Materia

- Clase 5: Codificación one-hot y dummy. Otros tipos de codificación (binaria, hashing, etc). Valores faltantes (missing values). Normalización de datos. Transformación de datos (box-cox transformation).
- Clase 6: Aplicación de Tests estadísticos univariados para Machine Learning. Ejemplos prácticos. Reducción de la dimensión. Normalización (Z-score). Principal Component Analysis (PCA).

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Programa de la Materia

- Clase 5: Codificación one-hot y dummy. Otros tipos de codificación (binaria, hashing, etc). Valores faltantes (missing values). Normalización de datos. Transformación de datos (box-cox transformation).
- Clase 6: Aplicación de Tests estadísticos univariados para Machine Learning. Ejemplos prácticos. Reducción de la dimensión. Normalización (Z-score). Principal Component Analysis (PCA).
- Clase 7: Taller práctico.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Programa de la Materia

- Clase 5: Codificación one-hot y dummy. Otros tipos de codificación (binaria, hashing, etc). Valores faltantes (missing values). Normalización de datos. Transformación de datos (box-cox transformation).
- Clase 6: Aplicación de Tests estadísticos univariados para Machine Learning. Ejemplos prácticos. Reducción de la dimensión. Normalización (Z-score). Principal Component Analysis (PCA).
- Clase 7: Taller práctico.
- Clase 8: Evaluación final.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua

FACULTAD
DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$
- Para variables discretas, $F(a) = \sum_{\forall x \leq a} P(X = x)$, donde $P(X = x)$ es la función de masa de probabilidad

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$
- Para variables discretas, $F(a) = \sum_{\forall x \leq a} P(X = x)$, donde $P(X = x)$ es la función de masa de probabilidad
- Para variables continuas, $F(x) = \int_{-\infty}^{x} p(u)du$, donde $p(\cdot)$ es la función de densidad de probabilidad (pdf, en inglés)

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$
- Para variables discretas, $F(a) = \sum_{\forall x \leq a} P(X = x)$, donde $P(X = x)$ es la función de masa de probabilidad
- Para variables continuas, $F(x) = \int_{-\infty}^{x} p(u)du$, donde $p(\cdot)$ es la función de densidad de probabilidad (pdf, en inglés)
- $F(\cdot)$ también es conocida como cdf (cumulative distribution function)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$
- Para variables discretas, $F(a) = \sum_{\forall x \leq a} P(X = x)$, donde $P(X = x)$ es la función de masa de probabilidad
- Para variables continuas, $F(x) = \int_{-\infty}^{x} p(u)du$, donde $p(\cdot)$ es la función de densidad de probabilidad (pdf, en inglés)
- $F(\cdot)$ también es conocida como cdf (cumulative distribution function)
- La pdf es la derivada de la cdf, $p(x) \triangleq dF(x)/dx$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$
- Para variables discretas, $F(a) = \sum_{\forall x \leq a} P(X = x)$, donde $P(X = x)$ es la función de masa de probabilidad
- Para variables continuas, $F(x) = \int_{-\infty}^{x} p(u)du$, donde $p(\cdot)$ es la función de densidad de probabilidad (pdf, en inglés)
- $F(\cdot)$ también es conocida como cdf (cumulative distribution function)
- La pdf es la derivada de la cdf, $p(x) \triangleq dF(x)/dx$
- $\int_{-\infty}^{+\infty} p(x)dx = 1$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución de Probabilidad

- La función de distribución de probabilidad $F(\cdot)$ de una variable aleatoria X para un número x es $F(x) = P(X \leq x)$
- Para v.a. discretas, $F(x)$ es no-negativa, y monótonamente creciente de a saltos
- Para v.a. continuas, $F(x)$ es no-negativa, monótonamente creciente y continua
- Para un intervalo $[a, b]$ tenemos que $P(a < X \leq b) = F(b) - F(a)$
- Para variables discretas, $F(a) = \sum_{\forall x \leq a} P(X = x)$, donde $P(X = x)$ es la función de masa de probabilidad
- Para variables continuas, $F(x) = \int_{-\infty}^{x} p(u)du$, donde $p(\cdot)$ es la función de densidad de probabilidad (pdf, en inglés)
- $F(\cdot)$ también es conocida como cdf (cumulative distribution function)
- La pdf es la derivada de la cdf, $p(x) \triangleq dF(x)/dx$
- $\int_{-\infty}^{+\infty} p(x)dx = 1$
- La cdf y la pdf son no-negativas

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Función de Distribución Conjunta y Marginal

- Es de interés conocer la relación entre dos o más variables

Función de Distribución Conjunta y Marginal

- Es de interés conocer la relación entre dos o más variables
- Se usa la distribución conjunta para tal fin

Función de Distribución Conjunta y Marginal

- Es de interés conocer la relación entre dos o más variables
- Se usa la distribución conjunta para tal fin
- La probabilidad conjunta de X, Y es $F(x, y) = P\{X \leq x, Y \leq y\}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución Conjunta y Marginal

- Es de interés conocer la relación entre dos o más variables
- Se usa la distribución conjunta para tal fin
- La probabilidad conjunta de X, Y es $F(x, y) = P\{X \leq x, Y \leq y\}$
- La distribución marginal es
  $F_X(x) = P(X \leq x) = P\{X \leq x, Y \leq \infty\} = F(x, \infty)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Es de interés conocer la relación entre dos o más variables
- Se usa la distribución conjunta para tal fin
- La probabilidad conjunta de X, Y es $F(x, y) = P\{X \leq x, Y \leq y\}$
- La distribución marginal es $F_X(x) = P(X \leq x) = P\{X \leq x, Y \leq \infty\} = F(x, \infty)$
- Caso discreto: $P(X \leq x) = \sum_j P(x, y_j)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Función de Distribución Conjunta y Marginal

- Es de interés conocer la relación entre dos o más variables
- Se usa la distribución conjunta para tal fin
- La probabilidad conjunta de X, Y es $F(x, y) = P\{X \leq x, Y \leq y\}$
- La distribución marginal es
  $F_X(x) = P(X \leq x) = P\{X \leq x, Y \leq \infty\} = F(x, \infty)$
- Caso discreto: $P(X \leq x) = \sum_j P(x, y_j)$
- Caso continuo: $p_X(x) = \int_{-\infty}^{+\infty} p(x, y)dy$

Distribuciones Condicionales

- La probabilidad condicional denota la probabilidad de ocurrencia de un evento X dado que Y ha ocurrido

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribuciones Condicionales

- La probabilidad condicional denota la probabilidad de ocurrencia de un evento X dado que Y ha ocurrido
- Caso discreto:
  $P_{X|Y} = P\{X = x|Y = y\} = \frac{P\{X=x, Y=y\}}{P(Y=y)} = \frac{P\{X=x, Y=y\}}{P_Y(y)}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribuciones Condicionales

- La probabilidad condicional denota la probabilidad de ocurrencia de un evento X dado que Y ha ocurrido
- Caso discreto: $P_{X|Y} = P\{X = x|Y = y\} = \frac{P\{X=x,Y=y\}}{P(Y=y)} = \frac{P\{X=x,Y=y\}}{P_Y(y)}$
- Caso continuo: $p_{X|Y} = \frac{p\{X=x,Y=y\}}{p_Y(y)}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribuciones Condicionales

- La probabilidad condicional denota la probabilidad de ocurrencia de un evento X dado que Y ha ocurrido
- Caso discreto:
  $P_{X|Y} = P\{X = x|Y = y\} = \frac{P\{X=x,Y=y\}}{P(Y=y)} = \frac{P\{X=x,Y=y\}}{P_Y(y)}$
- Caso continuo: $p_{X|Y} = \frac{p\{X=x,Y=y\}}{p_Y(y)}$
- ¿Qué sucede si X, Y son independientes?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribuciones Condicionales

- La probabilidad condicional denota la probabilidad de ocurrencia de un evento X dado que Y ha ocurrido
- Caso discreto:
  $P_{X|Y} = P\{X = x|Y = y\} = \frac{P\{X=x,Y=y\}}{P(Y=y)} = \frac{P\{X=x,Y=y\}}{P_Y(y)}$
- Caso continuo: $p_{X|Y} = \frac{p\{X=x,Y=y\}}{p_Y(y)}$
- ¿Qué sucede si X, Y son independientes?
- ¿Qué sucede si $X = Y$?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos
- Caso discreto: $\mu \triangleq E[X] = \sum_i x_i P(X = x_i)$

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos
- Caso discreto: $\mu \triangleq E[X] = \sum_i x_i P(X = x_i)$
- Caso continuo: $\mu \triangleq E[X] = \int_{-\infty}^{+\infty} xp(x)dx$

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos
- Caso discreto: $\mu \triangleq E[X] = \sum_i x_i P(X = x_i)$
- Caso continuo: $\mu \triangleq E[X] = \int_{-\infty}^{+\infty} x p(x) dx$
- La esperanza es un operador lineal: $E[aX + bY + c] = aE[X] + bE[Y] + c$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos
- Caso discreto: $\mu \triangleq E[X] = \sum_i x_i P(X = x_i)$
- Caso continuo: $\mu \triangleq E[X] = \int_{-\infty}^{+\infty} xp(x)dx$
- La esperanza es un operador lineal: $E[aX + bY + c] = aE[X] + bE[Y] + c$
- Funciones reales: $E[g(X)] = \sum_i g(x_i)P(X = x_i)$ or $E[g(x)] = \int g(x)p(x)dx$

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos
- Caso discreto: $\mu \triangleq E[X] = \sum_i x_i P(X = x_i)$
- Caso continuo: $\mu \triangleq E[X] = \int_{-\infty}^{+\infty} x p(x) dx$
- La esperanza es un operador lineal: $E[aX + bY + c] = aE[X] + bE[Y] + c$
- Funciones reales: $E[g(X)] = \sum_i g(x_i) P(X = x_i)$ or $E[g(x)] = \int g(x) p(x) dx$
- $E[X^n]$ es el n-esimo momento de X

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Esperanza

- La esperanza, valor esperado o media de una v.a. es el promedio aritmético en varios experimentos
- Caso discreto: $\mu \triangleq E[X] = \sum_i x_i P(X = x_i)$
- Caso continuo: $\mu \triangleq E[X] = \int_{-\infty}^{+\infty} xp(x)dx$
- La esperanza es un operador lineal: $E[aX + bY + c] = aE[X] + bE[Y] + c$
- Funciones reales: $E[g(X)] = \sum_i g(x_i)P(X = x_i)$ or $E[g(x)] = \int g(x)p(x)dx$
- $E[X^n]$ es el n-esimo momento de X
- La esperanza no siempre es el valor más probable

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Varianza

- Sea $\mu_X = E[X]$ la media de X, la varianza de X es $\mathrm{var}[X] = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2$

Varianza

- Sea $\mu_X = E[X]$ la media de X, la varianza de X es $\text{var}[X] = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2$
- Normalmente designaremos $\sigma_X^2 \triangleq \text{var}[X]$

Varianza

- Sea $\mu_X = E[X]$ la media de X, la varianza de X es $\text{var}[X] = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2$
- Normalmente designaremos $\sigma_X^2 \triangleq \text{var}[X]$
- $\text{var}[aX + b] = a^2\sigma_X^2$

Varianza

- Sea $\mu_X = E[X]$ la media de X, la varianza de X es $\text{var}[X] = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2$
- Normalmente designaremos $\sigma_X^2 \triangleq \text{var}[X]$
- $\text{var}[aX + b] = a^2\sigma_X^2$
- Por qué b desapareció de la expresión anterior?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Varianza

- Sea $\mu_X = E[X]$ la media de X, la varianza de X es $\text{var}[X] = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2$
- Normalmente designaremos $\sigma_X^2 \triangleq \text{var}[X]$
- $\text{var}[aX + b] = a^2\sigma_X^2$
- Por qué b desapareció de la expresión anterior?
- El desvío estándar es $\sigma_X = \sqrt{\sigma_X^2}$, medido en unidades de X

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Covarianza

- La covarianza indica la relación entre dos v.a.

Covarianza

- La covarianza indica la relación entre dos v.a.
- $\mathrm{cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Covarianza

- La covarianza indica la relación entre dos v.a.
- $\text{cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$
- $\text{cov}[X, X] = \text{var}[X]$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Covarianza

- La covarianza indica la relación entre dos v.a.
- $\text{cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$
- $\text{cov}[X, X] = \text{var}[X]$
- $\text{cov}[X + Y, Z] = \text{cov}[X, Z] + \text{cov}[Y, Z]$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Covarianza

- La covarianza indica la relación entre dos v.a.
- $\text{cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$
- $\text{cov}[X, X] = \text{var}[X]$
- $\text{cov}[X + Y, Z] = \text{cov}[X, Z] + \text{cov}[Y, Z]$
- $\text{var}[X + Y] = \text{var}[X] + \text{var}[Y] + 2\text{cov}[X, Y]$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Covarianza

- La covarianza indica la relación entre dos v.a.
- $\text{cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$
- $\text{cov}[X, X] = \text{var}[X]$
- $\text{cov}[X + Y, Z] = \text{cov}[X, Z] + \text{cov}[Y, Z]$
- $\text{var}[X + Y] = \text{var}[X] + \text{var}[Y] + 2\text{cov}[X, Y]$
- La correlación, o índice de correlación, se define como $\text{corr}[X, Y] = \frac{\text{cov}[X,Y]}{\sqrt{\text{var}[X]\text{var}[Y]}} \in [-1, +1]$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Covarianza

- La covarianza indica la relación entre dos v.a.
- $\mathrm{cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - \mu_X\mu_Y$
- $\mathrm{cov}[X,X] = \mathrm{var}[X]$
- $\mathrm{cov}[X+Y,Z] = \mathrm{cov}[X,Z] + \mathrm{cov}[Y,Z]$
- $\mathrm{var}[X+Y] = \mathrm{var}[X] + \mathrm{var}[Y] + 2\mathrm{cov}[X,Y]$
- La correlación, o índice de correlación, se define como $\mathrm{corr}[X,Y] = \frac{\mathrm{cov}[X,Y]}{\sqrt{\mathrm{var}[X]\mathrm{var}[Y]}} \in [-1,+1]$
- Normalmente usamos $\rho_{XY} \triangleq \mathrm{corr}[X,Y]$

Covarianza

- Si $\rho_{XY} > 0$, un cambio positivo en X influencia un cambio positivo en Y

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Covarianza

- Si $\rho_{XY} > 0$, un cambio positivo en X influencia un cambio positivo en Y
- Si $\rho_{XY} < 0$, un cambio positivo en X influencia un cambio negativo en Y

Covarianza

- Si $\rho_{XY} > 0$, un cambio positivo en X influencia un cambio positivo en Y
- Si $\rho_{XY} < 0$, un cambio positivo en X influencia un cambio negativo en Y
- Si $\rho_{XY} = 0$, un cambio en X no influencia un cambio en Y, se dice que las variables están descorrelacionadas

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Covarianza

- Si $\rho_{XY} > 0$, un cambio positivo en X influencia un cambio positivo en Y
- Si $\rho_{XY} < 0$, un cambio positivo en X influencia un cambio negativo en Y
- Si $\rho_{XY} = 0$, un cambio en X no influencia un cambio en Y, se dice que las variables están descorrelacionadas
- $\rho_{XY} = 0$ no implica que X, Y sean v.a. independientes (excepto si X, Y están normalmente distribuidas)

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Covarianza

- Si $\rho_{XY} > 0$, un cambio positivo en X influencia un cambio positivo en Y
- Si $\rho_{XY} < 0$, un cambio positivo en X influencia un cambio negativo en Y
- Si $\rho_{XY} = 0$, un cambio en X no influencia un cambio en Y, se dice que las variables están descorrelacionadas
- $\rho_{XY} = 0$ no implica que X, Y sean v.a. independientes (excepto si X, Y están normalmente distribuidas)
- Sin embargo, si X, Y son independientes, $\rho_{XY} = 0$

Momento de una V.A.

- Definimos el *n*-esimo momento de X como $E[X^n]$

Momento de una V.A.

- Definimos el n-esimo momento de X como $E[X^n]$
- Definimos la funcion generadora de momentos de X como:

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Momento de una V.A.

- Definimos el n-esimo momento de X como $E[X^n]$
- Definimos la funcion generadora de momentos de X como:
- $M_X(t) = E(e^{tX})$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Momento de una V.A.

- Definimos el n-esimo momento de X como $E[X^n]$
- Definimos la funcion generadora de momentos de X como:
- $M_X(t) = E(e^{tX})$
- Asumimos que la esperanza anterior $E(e^{tX})$ existe (i.e. es finita) para algun intervalo de $t \in (-\delta, +\delta)$ que contiene $t = 0$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Momento de una V.A.

- Definimos el n-esimo momento de X como $E[X^n]$
- Definimos la funcion generadora de momentos de X como:
- $M_X(t) = E(e^{tX})$
- Asumimos que la esperanza anterior $E(e^{tX})$ existe (i.e. es finita) para algun intervalo de $t \in (-\delta, +\delta)$ que contiene $t = 0$
- Para calcular $E[X^n]$, tomamos la derivada de $M_X(t)$ respecto de t y la evaluamos en $t = 0$, es decir:

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Momento de una V.A.

- Definimos el n-esimo momento de X como $E[X^n]$
- Definimos la funcion generadora de momentos de X como:
- $M_X(t) = E(e^{tX})$
- Asumimos que la esperanza anterior $E(e^{tX})$ existe (i.e. es finita) para algun intervalo de $t \in (-\delta, +\delta)$ que contiene $t = 0$
- Para calcular $E[X^n]$, tomamos la derivada de $M_X(t)$ respecto de t y la evaluamos en $t = 0$, es decir:
- $E[X^n] = \frac{d^n M_X(t)}{dt^n}(0) \triangleq M_X^{(n)}(0)$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Momento de una V.A. (Cont.)

- Encontrar $M_X(t)$ de una v.a. X con distribución exponencial y parámetro λ

Momento de una V.A. (Cont.)

- Encontrar $M_X(t)$ de una v.a. X con distribución exponencial y parámetro λ
- Usando las propiedades de esperanza de una función $g(X)$ de una v.a. X con distribución conocida, tenemos:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Momento de una V.A. (Cont.)

- Encontrar $M_X(t)$ de una v.a. X con distribución exponencial y parámetro λ
- Usando las propiedades de esperanza de una función $g(X)$ de una v.a. X con distribución conocida, tenemos:
- $M_X(t) = E[e^{tX}] = \int_0^{\infty} e^{tx}\, \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Momento de una V.A. (Cont.)

- Encontrar $M_X(t)$ de una v.a. X con distribución exponencial y parámetro λ
- Usando las propiedades de esperanza de una función $g(X)$ de una v.a. X con distribución conocida, tenemos:
- $M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx}\,\lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda$
- $E[X] = M_X^{(1)}(0) = \frac{1}{\lambda}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Momento de una V.A. (Cont.)

- Encontrar $M_X(t)$ de una v.a. X con distribución exponencial y parámetro λ
- Usando las propiedades de esperanza de una función $g(X)$ de una v.a. X con distribución conocida, tenemos:
- $M_X(t) = E[e^{tX}] = \int_0^{\infty} e^{tx}\,\lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda$
- $E[X] = M_X^{(1)}(0) = \frac{1}{\lambda}$
- $E[X^2] = M_X^{(2)}(0) = \frac{2}{\lambda^2}$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Momento de una V.A. (Cont.)

- Encontrar $M_X(t)$ de una v.a. X con distribución exponencial y parámetro λ
- Usando las propiedades de esperanza de una función $g(X)$ de una v.a. X con distribución conocida, tenemos:
- $M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx}\, \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad t < \lambda$
- $E[X] = M_X^{(1)}(0) = \frac{1}{\lambda}$
- $E[X^2] = M_X^{(2)}(0) = \frac{2}{\lambda^2}$
- $\text{var}[X] = E[X^2] - (E[X])^2 = \frac{1}{\lambda^2}$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Estadística de Orden k

- Sea $\boldsymbol{X}$ un vector aleatorio compuesto de N muestras, es decir : $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]^T$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Estadística de Orden k

- Sea $\boldsymbol{X}$ un vector aleatorio compuesto de N muestras, es decir : $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]^T$
- Definimos la estadística de orden k como el k-ésimo valor más pequeño de $\boldsymbol{X}$ luego de ordenarlo en forma ascendente

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Estadística de Orden k

- Sea $\boldsymbol{X}$ un vector aleatorio compuesto de N muestras, es decir : $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]^T$
- Definimos la estadística de orden k como el k-ésimo valor más pequeño de $\boldsymbol{X}$ luego de ordenarlo en forma ascendente
- Es decir, si $\boldsymbol{X_{ord}} = ordenarAscendente(\boldsymbol{X})$, entonces

Estadística de Orden k

- Sea $\boldsymbol{X}$ un vector aleatorio compuesto de N muestras, es decir : $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]^T$
- Definimos la estadística de orden k como el k-ésimo valor más pequeño de $\boldsymbol{X}$ luego de ordenarlo en forma ascendente
- Es decir, si $\boldsymbol{X_{ord}} = ordenarAscendente(\boldsymbol{X})$, entonces
- $\boldsymbol{X_{ord}}(k)$ es la estadística de orden k de $\boldsymbol{X}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Estadística de Orden k

- Sea $\boldsymbol{X}$ un vector aleatorio compuesto de N muestras, es decir : $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]^T$
- Definimos la estadística de orden k como el k-ésimo valor más pequeño de $\boldsymbol{X}$ luego de ordenarlo en forma ascendente
- Es decir, si $\boldsymbol{X_{ord}} = ordenarAscendente(\boldsymbol{X})$, entonces
- $\boldsymbol{X_{ord}}(k)$ es la estadística de orden k de $\boldsymbol{X}$
- ¿Qué representan $\boldsymbol{X_{ord}}(1)$ y $\boldsymbol{X_{ord}}(N)$?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Teorema Central del Límite

- Sean $X_1, X_2, \ldots, X_N$ un conjunto de v.a. i.i.d. cada una con media μ y varianza σ^2; para N grande se tiene que

- Sean $X_1, X_2, \ldots, X_N$ un conjunto de v.a. i.i.d. cada una con media μ y varianza σ^2; para N grande se tiene que
- $X_1 + X_2 + \ldots + X_N \sim \mathcal{N}(N\mu, N\sigma^2)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Teorema Central del Límite

- Sean $X_1, X_2, \ldots, X_N$ un conjunto de v.a. i.i.d. cada una con media μ y varianza σ^2; para N grande se tiene que
- $X_1 + X_2 + \ldots + X_N \sim \mathcal{N}(N\mu, N\sigma^2)$
- Es decir, la suma de N v.a. i.i.d. converge a una distribución normal con media y varianza igual a la suma de cada una

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Teorema Central del Límite

- Sean $X_1, X_2, \ldots, X_N$ un conjunto de v.a. i.i.d. cada una con media μ y varianza σ^2; para N grande se tiene que
- $X_1 + X_2 + \ldots + X_N \sim \mathcal{N}(N\mu, N\sigma^2)$
- Es decir, la suma de N v.a. i.i.d. converge a una distribución normal con media y varianza igual a la suma de cada una
- ¿Cuál será, aproximadamente, la distribución de $\frac{X_1+X_2+\ldots+X_N}{N}$?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$
- $P\{X = 1\} = p$ es la p. de éxito, $P\{X = 0\} = 1 - p$ la p. de fracaso

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$
- $P\{X = 1\} = p$ es la p. de éxito, $P\{X = 0\} = 1 - p$ la p. de fracaso
- $P\{X = i\} = p^i(1 - p)^i$, $i = \{0, 1\}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$
- $P\{X = 1\} = p$ es la p. de éxito, $P\{X = 0\} = 1 - p$ la p. de fracaso
- $P\{X = i\} = p^i(1 - p)^i$, $i = \{0, 1\}$
- $E[X] = p$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$
- $P\{X = 1\} = p$ es la p. de éxito, $P\{X = 0\} = 1 - p$ la p. de fracaso
- $P\{X = i\} = p^i(1 - p)^i$, $i = \{0, 1\}$
- $E[X] = p$
- $\text{var}[X] = p(1 - p)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$
- $P\{X = 1\} = p$ es la p. de éxito, $P\{X = 0\} = 1 - p$ la p. de fracaso
- $P\{X = i\} = p^i(1 - p)^i$, $i = \{0, 1\}$
- $E[X] = p$
- $\text{var}[X] = p(1 - p)$
- Binomial: N i.i.d. v.a. Bernoulli, X es la p. de i éxitos en N tiradas

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Bernoulli y Binomial

- Bernoulli: Ensayo de éxito (1) o fracaso (0)
- La v.a. del proceso toma los valores $X = \{0, 1\}$
- $P\{X = 1\} = p$ es la p. de éxito, $P\{X = 0\} = 1 - p$ la p. de fracaso
- $P\{X = i\} = p^i(1 - p)^i$, $i = \{0, 1\}$
- $E[X] = p$
- $\mathrm{var}[X] = p(1 - p)$
- Binomial: N i.i.d. v.a. Bernoulli, X es la p. de i éxitos en N tiradas
- $P\{X = i\} = \binom{N}{i} p^i (1 - p)^{N-i}$, $i = \{0, 1, \ldots, N\}$

Distribución Uniforme

- Una v.a. uniforme X en el intervalo cerrado $[a, b]$ tiene pdf:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Uniforme

- Una v.a. uniforme X en el intervalo cerrado $[a, b]$ tiene pdf:
- $p(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{otro } x \end{cases}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Uniforme

- Una v.a. uniforme X en el intervalo cerrado $[a, b]$ tiene pdf:
- $p(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{otro } x \end{cases}$
- $E[X] = \frac{a+b}{2}$

Distribución Uniforme

- Una v.a. uniforme X en el intervalo cerrado $[a, b]$ tiene pdf:
- $p(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{otro } x \end{cases}$
- $E[X] = \frac{a+b}{2}$
- $\text{var}[X] = \frac{(b-a)^2}{12}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \; -\infty < x < +\infty$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < +\infty$
- La normal cero-uno $\mathcal{Z} = \mathcal{N}(0, 1)$ es:

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < +\infty$
- La normal cero-uno $\mathcal{Z} = \mathcal{N}(0, 1)$ es:
- $p_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < +\infty$
- La normal cero-uno $\mathcal{Z} = \mathcal{N}(0, 1)$ es:
- $p_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$
- La cdf es $\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} p_Z(z)$

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < +\infty$
- La normal cero-uno $\mathcal{Z} = \mathcal{N}(0, 1)$ es:
- $p_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$
- La cdf es $\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} p_Z(z)$
- Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < +\infty$
- La normal cero-uno $\mathcal{Z} = \mathcal{N}(0, 1)$ es:
- $p_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$
- La cdf es $\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} p_Z(z)$
- Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$
- Normalización: Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Normal o Gaussiana

- Una v.a. normal o Gaussiana con media μ y varianza σ^2, denotada $\mathcal{N}(\mu, \sigma^2)$ tiene una pdf:
- $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < +\infty$
- La normal cero-uno $\mathcal{Z} = \mathcal{N}(0,1)$ es:
- $p_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$
- La cdf es $\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} p_Z(z)$
- Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$
- Normalización: Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$
- El teorema central del límite se puede aplicar aproximar la suma de v.a. i.i.d. como una variable normal

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Chi-Cuadrado

- Si Z_i son v.a. i.i.d $\mathcal{N}(0,1)$, la distribución de la suma de los cuadrados $X = Z_1^2 + Z_2^2 + \ldots + Z_n^2$ es:

Distribución Chi-Cuadrado

- Si Z_i son v.a. i.i.d $\mathcal{N}(0,1)$, la distribución de la suma de los cuadrados $X = Z_1^2 + Z_2^2 + \ldots + Z_n^2$ es:
- $X \sim \mathcal{X}_n^2$, donde $\mathcal{X}_n^2$ es una chi-cuadrado de orden n

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Chi-Cuadrado

- Si Z_i son v.a. i.i.d $\mathcal{N}(0, 1)$, la distribución de la suma de los cuadrados $X = Z_1^2 + Z_2^2 + \ldots + Z_n^2$ es:
- $X \sim \mathcal{X}_n^2$, donde $\mathcal{X}_n^2$ es una chi-cuadrado de orden n
- Usaremos la distribución chi-cuadrado para ensayar tests de hipótesis de varianza, y tests de bondad de ajuste (en inglés, goodness of fit) usando el teorema de Pearson

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución t de Student

- Si $Z \sim \mathcal{Z}$ y $X \sim \mathcal{X}_n^2$ son v.a. independientes entonces el cociente $T_n = \frac{Z}{\sqrt{X/n}}$ tiene distribución t con n grados de libertad

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución t de Student

- Si $Z \sim \mathcal{Z}$ y $X \sim \mathcal{X}_n^2$ son v.a. independientes entonces el cociente $T_n = \frac{Z}{\sqrt{X/n}}$ tiene distribución t con n grados de libertad
- $E[T_n] = 0, \; n > 1$

Distribución t de Student

- Si $Z \sim \mathcal{Z}$ y $X \sim \mathcal{X}_n^2$ son v.a. independientes entonces el cociente $T_n = \frac{Z}{\sqrt{X/n}}$ tiene distribución t con n grados de libertad
- $E[T_n] = 0,\ n > 1$
- $\mathrm{var}[T_n] = \frac{n}{n-2}$ para$n > 2$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Si $Z \sim \mathcal{Z}$ y $X \sim \mathcal{X}_n^2$ son v.a. independientes entonces el cociente $T_n = \frac{Z}{\sqrt{X/n}}$ tiene distribución t con n grados de libertad
- $E[T_n] = 0,\ n > 1$
- var$[T_n] = \frac{n}{n-2}$ para$n > 2$
- T_n se parece a $\mathcal{Z}$ cuando $n \to \infty$ pero con colas más notorias (mayor variabilidad que la normal cero-uno)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución t de Student

- Si $Z \sim \mathcal{Z}$ y $X \sim \mathcal{X}_n^2$ son v.a. independientes entonces el cociente $T_n = \frac{Z}{\sqrt{X/n}}$ tiene distribución t con n grados de libertad
- $E[T_n] = 0,\ n > 1$
- $\mathrm{var}[T_n] = \frac{n}{n-2}$ para$n > 2$
- T_n se parece a $\mathcal{Z}$ cuando $n \to \infty$ pero con colas más notorias (mayor variabilidad que la normal cero-uno)
- Usaremos la distribución t de Student para ensayar un test de hipótesis y estimación de intervalo de confianza

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i
- La media muestral es:

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i
- La media muestral es:
- $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i
- La media muestral es:
- $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$
- La varianza muestral es:

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i
- La media muestral es:
- $\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} X_i$
- La varianza muestral es:
- $s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \hat{\mu})^2$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i
- La media muestral es:
- $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$
- La varianza muestral es:
- $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \hat{\mu})^2$
- El desvío estándar o dispersión muestral es:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Media y Varianza Muestral

- Sea X una v.a. con N muestras denominadas X_i
- La media muestral es:
- $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i$
- La varianza muestral es:
- $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \hat{\mu})^2$
- El desvío estándar o dispersión muestral es:
- $s = \sqrt{s^2}$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Media Recortada

- La media es el centro de gravedad de la distribución

Media Recortada

- La media es el centro de gravedad de la distribución
- Como ya conocemos, la media muestral la calculamos como el promedio aritmético de las N muestras

Media Recortada

- La media es el centro de gravedad de la distribución
- Como ya conocemos, la media muestral la calculamos como el promedio aritmético de las N muestras
- Existen otras *medidas de tendencia central* de interés para el análisis de datos

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Media Recortada

- La media es el centro de gravedad de la distribución
- Como ya conocemos, la media muestral la calculamos como el promedio aritmético de las N muestras
- Existen otras *medidas de tendencia central* de interés para el análisis de datos
- Cuando la distribución tiene valores extremos, conviene recortar la distribución para analizar los datos

Media Recortada

- La media es el centro de gravedad de la distribución
- Como ya conocemos, la media muestral la calculamos como el promedio aritmético de las N muestras
- Existen otras *medidas de tendencia central* de interés para el análisis de datos
- Cuando la distribución tiene valores extremos, conviene recortar la distribución para analizar los datos
- La media recortada a un $p\%$ de X es la media que se obtiene de la distribución al descartar el $p\%$ superior y el $p\%$ inferior

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Media Recortada

- La media es el centro de gravedad de la distribución
- Como ya conocemos, la media muestral la calculamos como el promedio aritmétrico de las N muestras
- Existen otras *medidas de tendencia central* de interés para el análisis de datos
- Cuando la distribución tiene valores extremos, conviene recortar la distribución para analizar los datos
- La media recortada a un $p\%$ de X es la media que se obtiene de la distribución al descartar el $p\%$ superior y el $p\%$ inferior
- Ejemplo

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Mediana

- La mediana es el valor que parte la distribución en dos porciones de área iguales a $(1/2)$

Mediana

- La mediana es el valor que parte la distribución en dos porciones de área iguales a (1/2)
- La mediana también se puede definir como la media recortada de 50%

Mediana

- La mediana es el valor que parte la distribución en dos porciones de área iguales a (1/2)
- La mediana también se puede definir como la media recortada de 50%
- La moda es el valor más probable (con mayor ocurrencia) de la distribución - recuerden máxima verosimilitud!

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Mediana

- La mediana es el valor que parte la distribución en dos porciones de área iguales a (1/2)
- La mediana también se puede definir como la media recortada de 50%
- La moda es el valor más probable (con mayor ocurrencia) de la distribución - recuerden máxima verosimilitud!
- Puede haber más de una moda

FACULTAD
DE INGENIERIA
Universidad de Buenos Aires

Mediana

- La mediana es el valor que parte la distribución en dos porciones de área iguales a (1/2)
- La mediana también se puede definir como la media recortada de 50%
- La moda es el valor más probable (con mayor ocurrencia) de la distribución - recuerden máxima verosimilitud!
- Puede haber más de una moda
- Ejemplo

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Oblicuidad, Asimetría o Skewness

- La media, la moda, la mediana y la media recortada son valores de tendencia central de una distribución

Oblicuidad, Asimetría o Skewness

- La media, la moda, la mediana y la media recortada son valores de tendencia central de una distribución
- La oblicuidad se puede definir como una medida de la falta de simetría

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Oblicuidad, Asimetría o Skewness

- La media, la moda, la mediana y la media recortada son valores de tendencia central de una distribución
- La oblicuidad se puede definir como una medida de la falta de simetría
- La oblicuidad de la distribución de una v.a. X con muestras X_i, media μ y varianza σ^2 se define matemáticamente como el tercer momento central estandarizado (i.e. dividido por la dispersión al cubo)

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Oblicuidad, Asimetría o Skewness

- La media, la moda, la mediana y la media recortada son valores de tendencia central de una distribución
- La oblicuidad se puede definir como una medida de la falta de simetría
- La oblicuidad de la distribución de una v.a. X con muestras X_i, media μ y varianza σ^2 se define matemáticamente como el tercer momento central estandarizado (i.e. dividido por la dispersión al cubo)
- $\tilde{\mu}_3 = E[\left(\frac{X-\mu}{\sigma}\right)^3] = \frac{\mu_3}{\sigma^3}$

Oblicuidad, Asimetría o Skewness

- La media, la moda, la mediana y la media recortada son valores de tendencia central de una distribución
- La oblicuidad se puede definir como una medida de la falta de simetría
- La oblicuidad de la distribución de una v.a. X con muestras X_i, media μ y varianza σ^2 se define matemáticamente como el tercer momento central estandarizado (i.e. dividido por la dispersión al cubo)
- $\tilde{\mu}_3 = E[\left(\frac{X-\mu}{\sigma}\right)^3] = \frac{\mu_3}{\sigma^3}$
- En distribuciones simétricas respecto de un pico, la moda, mediana,media y media recortada coinciden

Oblicuidad, Asimetría o Skewness

- La media, la moda, la mediana y la media recortada son valores de tendencia central de una distribución
- La oblicuidad se puede definir como una medida de la falta de simetría
- La oblicuidad de la distribución de una v.a. X con muestras X_i, media μ y varianza σ^2 se define matemáticamente como el tercer momento central estandarizado (i.e. dividido por la dispersión al cubo)
- $\tilde{\mu}_3 = E[\left(\frac{X-\mu}{\sigma}\right)^3] = \frac{\mu_3}{\sigma^3}$
- En distribuciones simétricas respecto de un pico, la moda, mediana,media y media recortada coinciden
- En este caso la oblicuidad es cero

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Skewness (Cont.)

- Si la distribución tiene más área hacia la izquierda de la moda que hacia la derecha, está oblicuada hacia la izquierda (negativa)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Skewness (Cont.)

- Si la distribución tiene más área hacia la izquierda de la moda que hacia la derecha, está oblicuada hacia la izquierda (negativa)
- Si la distribución tiene más área hacia la derecha de la moda que hacia la izquierda, está oblicuada hacia la derecha (positiva)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Skewness (Cont.)

- Si la distribución tiene más área hacia la izquierda de la moda que hacia la derecha, está oblicuada hacia la izquierda (negativa)
- Si la distribución tiene más área hacia la derecha de la moda que hacia la izquierda, está oblicuada hacia la derecha (positiva)
- Ejemplo

Skewness (Cont.)

- Si la distribución tiene más área hacia la izquierda de la moda que hacia la derecha, está oblicuada hacia la izquierda (negativa)
- Si la distribución tiene más área hacia la derecha de la moda que hacia la izquierda, está oblicuada hacia la derecha (positiva)
- Ejemplo
- La media sigue la dirección de la oblicuidad, respecto de la moda

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Skewness (Cont.)

- Si la distribución tiene más área hacia la izquierda de la moda que hacia la derecha, está oblicuada hacia la izquierda (negativa)
- Si la distribución tiene más área hacia la derecha de la moda que hacia la izquierda, está oblicuada hacia la derecha (positiva)
- Ejemplo
- La media sigue la dirección de la oblicuidad, respecto de la moda
- Entre la media y la moda se encuentra la mediana

Skewness (Cont.)

- Si la distribución tiene más área hacia la izquierda de la moda que hacia la derecha, está oblicuada hacia la izquierda (negativa)
- Si la distribución tiene más área hacia la derecha de la moda que hacia la izquierda, está oblicuada hacia la derecha (positiva)
- Ejemplo
- La media sigue la dirección de la oblicuidad, respecto de la moda
- Entre la media y la moda se encuentra la mediana
- Ejemplo

Curtosis

- La curtosis es una medida de cuán punteaguda o achatada es una distribución

Curtosis

- La curtosis es una medida de cuán punteaguda o achatada es una distribución
- La curtosis se calcula como el cuarto momento central estandarizado (i.e. dividido por la dispersión a la cuarta):

Curtosis

- La curtosis es una medida de cuán punteaguda o achatada es una distribución
- La curtosis se calcula como el cuarto momento central estandarizado (i.e. dividido por la dispersión a la cuarta):
- $\tilde{\mu}_4 = E[\left(\frac{X-\mu}{\sigma}\right)^4] = \frac{\mu_4}{\sigma^4}$

Curtosis

- La curtosis es una medida de cuán punteaguda o achatada es una distribución
- La curtosis se calcula como el cuarto momento central estandarizado (i.e. dividido por la dispersión a la cuarta):
- $\tilde{\mu}_4 = E[\left(\frac{X-\mu}{\sigma}\right)^4] = \frac{\mu_4}{\sigma^4}$
- La curtosis de una distribución normal es 3

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Curtosis

- La curtosis es una medida de cuán punteaguda o achatada es una distribución
- La curtosis se calcula como el cuarto momento central estandarizado (i.e. dividido por la dispersión a la cuarta):
- $\tilde{\mu}_4 = E[\left(\frac{X-\mu}{\sigma}\right)^4] = \frac{\mu_4}{\sigma^4}$
- La curtosis de una distribución normal es 3
- Se define por esto la *curtosis en exceso* de una distribución como $\tilde{\mu}_4 - 3$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Percentiles

- Las realizaciones de una v.a. se alejan de la media en función de la *variabilidad* de los datos

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Percentiles

- Las realizaciones de una v.a. se alejan de la media en función de la *variabilidad* de los datos
- Se define el *rango* de un conjunto de medidas como la diferencia entre el valor más grande y el más chico del conjunto

Percentiles

- Las realizaciones de una v.a. se alejan de la media en función de la *variabilidad* de los datos
- Se define el *rango* de un conjunto de medidas como la diferencia entre el valor más grande y el más chico del conjunto
- El rango por sí solo no aporta información precisa de la variabilidad

FACULTAD
DE INGENIERIA
Universidad de Buenos Aires

Percentiles

- Las realizaciones de una v.a. se alejan de la media en función de la *variabilidad* de los datos
- Se define el *rango* de un conjunto de medidas como la diferencia entre el valor más grande y el más chico del conjunto
- El rango por sí solo no aporta información precisa de la variabilidad
- Se define el percentil p-ésimo como aquel valor x_p en un conjunto *ordenado* de datos para el cual el $p\%$ de las mediciones están por debajo de x_p, y $(1-p)\%$ de las mediciones están por encima de x_p

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Para encontrar el percentil $i\%$ en base a N observaciones de X, denominadas X_i hacemos lo siguiente:

Percentiles (Cont.)

- Para encontrar el percentil $i\%$ en base a N observaciones de X, denominadas X_i hacemos lo siguiente:
- Primero ordenamos en forma ascendente las muestras de manera que tenemos $X_1 \leq X_2 \leq \ldots \leq X_i \leq \ldots \leq X_N$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Percentiles (Cont.)

- Para encontrar el percentil $i\%$ en base a N observaciones de X, denominadas X_i hacemos lo siguiente:
- Primero ordenamos en forma ascendente las muestras de manera que tenemos $X_1 \leq X_2 \leq \ldots \leq X_i \leq \ldots \leq X_N$
- El valor X_i corresponde al percentil $x_i = 100(i - 0.5)/N$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Percentiles (Cont.)

- Para encontrar el percentil $i\%$ en base a N observaciones de X, denominadas X_i hacemos lo siguiente:
- Primero ordenamos en forma ascendente las muestras de manera que tenemos $X_1 \leq X_2 \leq \ldots \leq X_i \leq \ldots \leq X_N$
- El valor X_i corresponde al percentil $x_i = 100(i - 0.5)/N$
- En la ecuación anterior, 0.5 para no asignar un percentil 100% a la muestra más grande (improbable)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Ciertos percentiles específicos son de interés, estos son:

Cuantiles y Cuartiles

- Ciertos percentiles específicos son de interés, estos son:
- El percentil $25\%, 50\%, 75\%$

Cuantiles y Cuartiles

- Ciertos percentiles específicos son de interés, estos son:
- El percentil $25\%, 50\%, 75\%$
- Estos también se denominan cuartil inferior $Q_1 = Q(.25)$, medio $Q_2 = Q(.5)$ y superior $Q_3 = Q(.75)$, respectivamente

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Cuantiles y Cuartiles

- Ciertos percentiles específicos son de interés, estos son:
- El percentil $25\%, 50\%, 75\%$
- Estos también se denominan cuartil inferior $Q_1 = Q(.25)$, medio $Q_2 = Q(.5)$ y superior $Q_3 = Q(.75)$, respectivamente
- ¿Cómo se llama al percentil 50%?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Cuantiles y Cuartiles

- Ciertos percentiles específicos son de interés, estos son:
- El percentil $25\%, 50\%, 75\%$
- Estos también se denominan cuartil inferior $Q_1 = Q(.25)$, medio $Q_2 = Q(.5)$ y superior $Q_3 = Q(.75)$, respectivamente
- ¿Cómo se llama al percentil 50%?
- Al mismo tiempo, se define como primero, segundo, ..., q-ésimo cuantil q a aquella subdivisión de la distribución en q partes iguales

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Cuantiles y Cuartiles

- Ciertos percentiles específicos son de interés, estos son:
- El percentil $25\%, 50\%, 75\%$
- Estos también se denominan cuartil inferior $Q_1 = Q(.25)$, medio $Q_2 = Q(.5)$ y superior $Q_3 = Q(.75)$, respectivamente
- ¿Cómo se llama al percentil 50%?
- Al mismo tiempo, se define como primero, segundo, ..., q-ésimo cuantil q a aquella subdivisión de la distribución en q partes iguales
- Los cuartiles son un caso particular de los cuantiles ($q = 4$)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Rango Intercuartil (IQR)

- El rango intercuartil (IQR) se define como la diferencia entre el cuartil superior y el inferior

Rango Intercuartil (IQR)

- El rango intercuartil (IQR) se define como la diferencia entre el cuartil superior y el inferior
- $IQR = Q_3 - Q_1$

Rango Intercuartil (IQR)

- El rango intercuartil (IQR) se define como la diferencia entre el cuartil superior y el inferior
- $IQR = Q_3 - Q_1$
- IQR mide la variabilidad en la región central de la distribución, pero ignora la variabilidad en los extremos superiores ($> 75\%$) e inferiores ($< 25\%$)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Rango Intercuartil (IQR)

- El rango intercuartil (IQR) se define como la diferencia entre el cuartil superior y el inferior
- $IQR = Q_3 - Q_1$
- IQR mide la variabilidad en la región central de la distribución, pero ignora la variabilidad en los extremos superiores ($> 75\%$) e inferiores ($< 25\%$)
- Ejemplo: calcular el IQR para la siguiente muestra de datos: $(20, 50, 50, 50, 50, 50, 50, 50, 50, 80)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Rango Intercuartil (IQR)

- El rango intercuartil (IQR) se define como la diferencia entre el cuartil superior y el inferior
- $IQR = Q_3 - Q_1$
- IQR mide la variabilidad en la región central de la distribución, pero ignora la variabilidad en los extremos superiores ($> 75\%$) e inferiores ($< 25\%$)
- Ejemplo: calcular el IQR para la siguiente muestra de datos: $(20, 50, 50, 50, 50, 50, 50, 50, 50, 80)$
- Notar que IQR es cero porque la media, la mediana y los cuartiles superiores e inferiores dan todos 50

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Rango Intercuartil (IQR)

- El rango intercuartil (IQR) se define como la diferencia entre el cuartil superior y el inferior
- $IQR = Q_3 - Q_1$
- IQR mide la variabilidad en la región central de la distribución, pero ignora la variabilidad en los extremos superiores ($> 75\%$) e inferiores ($< 25\%$)
- Ejemplo: calcular el IQR para la siguiente muestra de datos: $(20, 50, 50, 50, 50, 50, 50, 50, 50, 80)$
- Notar que IQR es cero porque la media, la mediana y los cuartiles superiores e inferiores dan todos 50
- Sin embargo los datos tienen variabilidad

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

IQR (Cont.)

- Por lo general es necesario contar con la siguiente mínima información para describir un conjunto de datos ordenado:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

IQR (Cont.)

- Por lo general es necesario contar con la siguiente mínima información para describir un conjunto de datos ordenado:
- El valor mínimo de la muestra (X_1)

IQR (Cont.)

- Por lo general es necesario contar con la siguiente mínima información para describir un conjunto de datos ordenado:
- El valor mínimo de la muestra (X_1)
- El cuartil inferior Q_1

IQR (Cont.)

- Por lo general es necesario contar con la siguiente mínima información para describir un conjunto de datos ordenado:
- El valor mínimo de la muestra (X_1)
- El cuartil inferior Q_1
- El cuartil medio Q_2 (la mediana)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

IQR (Cont.)

- Por lo general es necesario contar con la siguiente mínima información para describir un conjunto de datos ordenado:
- El valor mínimo de la muestra (X_1)
- El cuartil inferior Q_1
- El cuartil medio Q_2 (la mediana)
- El cuartil superior Q_3

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

IQR (Cont.)

- Por lo general es necesario contar con la siguiente mínima información para describir un conjunto de datos ordenado:
- El valor mínimo de la muestra (X_1)
- El cuartil inferior Q_1
- El cuartil medio Q_2 (la mediana)
- El cuartil superior Q_3
- El valor máximo de la muestra (X_N)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Regla Empírica

- Sea $\hat{\mu}$ la media muestral y s el desvío estándar muestral

Regla Empírica

- Sea $\hat{\mu}$ la media muestral y s el desvío estándar muestral
- La Regla Empírica dice que para una distribución con forma de montaña, se cumple:

Regla Empírica

- Sea $\hat{\mu}$ la media muestral y s el desvío estándar muestral
- La Regla Empírica dice que para una distribución con forma de montaña, se cumple:
- El intervalo $\hat{\mu} \pm s$ contiene el 68% de las mediciones

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Regla Empírica

- Sea $\hat{\mu}$ la media muestral y s el desvío estándar muestral
- La Regla Empírica dice que para una distribución con forma de montaña, se cumple:
- El intervalo $\hat{\mu} \pm s$ contiene el 68% de las mediciones
- El intervalo $\hat{\mu} \pm 2s$ contiene el 95% de las mediciones

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Regla Empírica

- Sea $\hat{\mu}$ la media muestral y s el desvío estándar muestral
- La Regla Empírica dice que para una distribución con forma de montaña, se cumple:
- El intervalo $\hat{\mu} \pm s$ contiene el 68% de las mediciones
- El intervalo $\hat{\mu} \pm 2s$ contiene el 95% de las mediciones
- El intervalo $\hat{\mu} \pm 3s$ contiene el 99.7% de las mediciones

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Estimación de Desvío Estándar con Rango

- Dado que el intervalo $\hat{\mu} \pm 2s$ contiene el 95% de las mediciones,...

Estimación de Desvío Estándar con Rango

- Dado que el intervalo $\hat{\mu} \pm 2s$ contiene el 95% de las mediciones,...
- el rango (máximo - mínimo) de las mediciones tiene longitud aproximada $2s - (-2s) = 4s$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Dado que el intervalo $\hat{\mu} \pm 2s$ contiene el 95% de las mediciones,...
- el rango (máximo - mínimo) de las mediciones tiene longitud aproximada $2s - (-2s) = 4s$
- Entonces podemos estimar s como:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Estimación de Desvío Estándar con Rango

- Dado que el intervalo $\hat{\mu} \pm 2s$ contiene el 95% de las mediciones,...
- el rango (máximo - mínimo) de las mediciones tiene longitud aproximada $2s - (-2s) = 4s$
- Entonces podemos estimar s como:
- $\hat{s} = \frac{rango}{4}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Coeficiente de Variación

- Una medida de cuanto varían las muestras respecto de la media es encontrar el coeficiente de variación (CV)

Coeficiente de Variación

- Una medida de cuanto varían las muestras respecto de la media es encontrar el coeficiente de variación (CV)
- Definimos el CV como:

Coeficiente de Variación

- Una medida de cuanto varían las muestras respecto de la media es encontrar el coeficiente de variación (CV)
- Definimos el CV como:
- $CV(\%) = \frac{\sigma}{|\mu|} \times 100$, para $\mu \neq 0$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Coeficiente de Variación

- Una medida de cuanto varían las muestras respecto de la media es encontrar el coeficiente de variación (CV)
- Definimos el CV como:
- $CV(\%) = \frac{\sigma}{|\mu|} \times 100$, para $\mu \neq 0$
- Esto sirve para comparar el grado de variación de distintos procesos estocásticos, normalizado por la media

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Coeficiente de Variación

- Una medida de cuanto varían las muestras respecto de la media es encontrar el coeficiente de variación (CV)
- Definimos el CV como:
- $CV(\%) = \frac{\sigma}{|\mu|} \times 100$, para $\mu \neq 0$
- Esto sirve para comparar el grado de variación de distintos procesos estocásticos, normalizado por la media
- Usando datos muestrales, el CV se estima como:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Coeficiente de Variación

- Una medida de cuanto varían las muestras respecto de la media es encontrar el coeficiente de variación (CV)
- Definimos el CV como:
- $CV(\%) = \frac{\sigma}{|\mu|} \times 100$, para $\mu \neq 0$
- Esto sirve para comparar el grado de variación de distintos procesos estocásticos, normalizado por la media
- Usando datos muestrales, el CV se estima como:
- $CV(\%) = \frac{s}{|\hat{\mu}|} \times 100$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:
- Identificamos Q_1, Q_2, Q_3

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:
- Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:
- Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:
- Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta el valor máximo de las muestras (Whisker)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:
- Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta el valor máximo de las muestras (Whisker)
- Trazamos una línea desde el lado Q_1 hasta el valor mínimo de las muestras (Whisker)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Rápida

- Para representar gráficamente la distribución de los datos en forma más compacta que en un histograma, se usan los denonimados gráficos de Box-and-Whiskers (B& W)
- Para graficarlos, hacemos lo siguiente:
- Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta el valor máximo de las muestras (Whisker)
- Trazamos una línea desde el lado Q_1 hasta el valor mínimo de las muestras (Whisker)
- Ejemplo

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta valor adyacente superior

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta valor adyacente superior
- Trazamos una línea desde el lado Q_1 hasta valor adyacente inferior

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta valor adyacente superior
- Trazamos una línea desde el lado Q_1 hasta valor adyacente inferior
- Marcamos los outliers ligeros (entre $1.5 \sim 3 \times IQR$) con una cruz, y los outliers pesados ($> 3 \times IQR$) con un círculo

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Box-and-Whiskers - Versión Completa

- Para hacer un gráfico de B& W completo debemos definir primero definimos lo siguiente:
- Definimos el *paso* como $paso = 1.5 \times IQR$:
- Definimos a la muestra x como un *outlier* si cumple lo siguiente:
- $x > Q_3 + paso$, o $x < Q_1 - paso$
- Definimos *valores adyacentes* a los últimos valores que no son outliers
- Para graficar Identificamos Q_1, Q_2, Q_3
- Graficamos una caja (Box) rectangular entre Q_1 y Q_3
- Trazamos una línea paralela los lados más corto de la caja Q_2
- Trazamos una línea desde el lado Q_3 hasta valor adyacente superior
- Trazamos una línea desde el lado Q_1 hasta valor adyacente inferior
- Marcamos los outliers ligeros (entre $1.5 \sim 3 \times IQR$) con una cruz, y los outliers pesados ($> 3 \times IQR$) con un círculo
- Ejemplo

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)
- Para esto, mostramos los dígitos *tallo*, es decir el o los primeros dígitos de un número excepto el último dígito

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)
- Para esto, mostramos los dígitos *tallo*, es decir el o los primeros dígitos de un número excepto el último dígito
- El tallo puede tener cualquier cantidad de números

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)
- Para esto, mostramos los dígitos *tallo*, es decir el o los primeros dígitos de un número excepto el último dígito
- El tallo puede tener cualquier cantidad de números
- Luego trazamos una línea vertical y trazamos el último dígito de cada valor de la v.a.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)
- Para esto, mostramos los dígitos *tallo*, es decir el o los primeros dígitos de un número excepto el último dígito
- El tallo puede tener cualquier cantidad de números
- Luego trazamos una línea vertical y trazamos el último dígito de cada valor de la v.a.
- Separamos el último dígito de cada valor arrojado por la v.a. con un espacio

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)
- Para esto, mostramos los dígitos *tallo*, es decir el o los primeros dígitos de un número excepto el último dígito
- El tallo puede tener cualquier cantidad de números
- Luego trazamos una línea vertical y trazamos el último dígito de cada valor de la v.a.
- Separamos el último dígito de cada valor arrojado por la v.a. con un espacio
- Cada hoja solo puede tener un número

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf

- Si queremos visualizar *todas* las observaciones de una realización de una v.a. en forma *agrupada* podemos usar los diagramas de Stem-and-Leaf (tallo y hoja)
- Para esto, mostramos los dígitos *tallo*, es decir el o los primeros dígitos de un número excepto el último dígito
- El tallo puede tener cualquier cantidad de números
- Luego trazamos una línea vertical y trazamos el último dígito de cada valor de la v.a.
- Separamos el último dígito de cada valor arrojado por la v.a. con un espacio
- Cada hoja solo puede tener un número
- Si no queremos mostrar valores decimales, redondeamos al entero más cercano

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf) (Cont.)

- Nota: si rotamos 90 grados en sentido antihorario al diagrama Stem-and-Leaf obtenemos una especie de *distribución*

Diagramas de Stem-and-Leaf) (Cont.)

- Nota: si rotamos 90 grados en sentido antihorario al diagrama Stem-and-Leaf obtenemos una especie de *distribución*
- Ejemplo: sea una v.a. X con observaciones $\{1.2 \quad 1.3 \quad 1.5 \quad 1.9 \quad 2.0 \quad 2.4 \quad 2.5 \quad 3.5 \quad 3.6 \quad 4.1\}$. Graficar el diagrama de stem-and-leaf

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Diagramas de Stem-and-Leaf) (Cont.)

- Nota: si rotamos 90 grados en sentido antihorario al diagrama Stem-and-Leaf obtenemos una especie de *distribución*
- Ejemplo: sea una v.a. X con observaciones $\{1.2 \quad 1.3 \quad 1.5 \quad 1.9 \quad 2.0 \quad 2.4 \quad 2.5 \quad 3.5 \quad 3.6 \quad 4.1\}$. Graficar el diagrama de stem-and-leaf
- El diagrama es:

Stem	Leaf
1	2 3 5 9
2	0 4 5
3	5 6
4	1

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicio 1

Simular $N = 100$ muestras de una v.a. X con distribución normal cero-uno.

a. Encontrar la curtosis y oblicuidad usando una rutina de software
b. Encontrar el IQR
c. Estimar el desvío estándar usando el rango de las mediciones simuladas
d. Trazar el diagrama de Box-and-Whiskers para una realización

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicio 2

Simular $N = 100$ muestras de una v.a. X con distribución exponencial con parámetro $\lambda = 1$.

a. Encontrar la curtosis y oblicuidad usando una rutina de software
b. Encontrar el IQR
c. Estimar el desvío estándar usando el rango de las mediciones simuladas
d. Trazar el diagrama de Box-and-Whiskers para una realización

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicio 3

Quiero saber si me conviene comprar acciones de Amazon, basado en la información financiera del Nasdaq de los últimos seis meses. Sólo me interesa comprarlas si voy a tener una probabilidad de al menos 95% de generar una ganancia mínima (promedio) de 2% si vendo las acciones durante el mes posterior a la compra. Usando las herramientas vistas en la clase de hoy y en Probabilidad y Estadística, ¿me recomiendan comprar las acciones o no?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Bibliografía I

E. Alpaydin, *Introduction to machine learning*.
MIT press, 2020.

R. L. Ott and M. T. Longnecker, *An introduction to statistical methods and data analysis*.
Nelson Education, 2015.

"Stony Brook University: Notes on Expectation, Moment Generating Functions, Variance, Covariance." `http://www.ams.sunysb.edu/~jsbm/courses/311/expectations.pdf`.
Accessed: 2021-02-07.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires