

Colección de Ejercicios Propuestos y sus resoluciones

Sebastian Elizalde
Padrón Nro. 96092
sebi.elizalde@gmail.com

Julian Ferres
Padrón Nro. 101483
julianferres@gmail.com

2do. Cuatrimestre de 2018

Aprendizaje Estadístico - Sebastian Grymberg
Facultad de Ingeniería, Universidad de Buenos Aires

Contents

1	Introducción	4
2	Lista de enunciados	4
2.1	Clase 1	4
2.1.1	Ejercicio 1	4
2.1.2	Ejercicio 2: Esperanza condicional	4
2.1.3	Ejercicio 3 (STOP) (SIMULACIÓN)	4
2.1.4	Ejercicio 4 (no entregable)	5
2.2	Clase 2	5
2.2.1	Ejercicio 5	5
2.2.2	Ejercicio 6	5
2.3	Clase 3	5
2.3.1	Ejercicio 7	5
2.3.2	Ejercicio 8 (STOP)(SIMULACIÓN)	6
2.4	Clase 4	6
2.4.1	Ejercicio 9: KNN	6
2.4.2	Ejercicio 9 bis	7
2.4.3	Ejercicio 10	7
2.5	Clase 5	7
2.5.1	Ejercicio 11	7
2.5.2	Ejercicio 12	7
2.5.3	Ejercicio 13	7
2.5.4	Ejercicio 14 (STOP)	8
2.6	Clase 6	8
2.6.1	Ejercicio 15	8
2.6.2	Ejercicio 16	8
3	Soluciones	9
3.1	Clase 1	9
3.1.1	Ejercicio 1	9
3.1.2	Ejercicio 2: Esperanza condicional	9
3.1.3	Ejercicio 3 (STOP) (SIMULACIÓN)	10
3.1.4	Ejercicio 4 (no entregable)	10
3.2	Clase 2	10
3.2.1	Ejercicio 5	10
3.2.2	Ejercicio 6	10
3.3	Clase 3	11
3.3.1	Ejercicio 7	11
3.3.2	Ejercicio 8 (STOP)(SIMULACIÓN)	11
3.4	Clase 4	11
3.4.1	Ejercicio 9: KNN	11
3.4.2	Ejercicio 9 bis	12
3.4.3	Ejercicio 10	13
3.5	Clase 5	13
3.5.1	Ejercicio 11	13
3.5.2	Ejercicio 12	13
3.5.3	Ejercicio 13	14
3.5.4	Ejercicio 14 (STOP)	14

3.6	Clase 6	14
3.6.1	Ejercicio 15	14
3.6.2	Ejercicio 16	14

1 Introducción

En este informe se busca recopilar los diferentes ejercicios propuestos en la asignatura *Aprendizaje Estadístico* en el 2^{do} cuatrimestre de 2018 por el profesor Sebastian Grymberg, junto con las soluciones propuestas por los integrantes del grupo.

2 Lista de enunciados

2.1 Clase 1

2.1.1 Ejercicio 1

Sea una función convexa $f(x)$ no negativa, x_0, x_1, x_2 tales que:

- $x_1 \leq x_0 \leq x_2$
- $x_0 = p_1 x_1 + p_2 x_2$
- $p_1 \geq 0, p_2 \geq 0, p_1 + p_2 = 1$

Y sea una recta g tal que:

- $g(x_1) = f(x_1)$
- $g(x_2) = f(x_2)$

Comprobar la siguiente desigualdad:

$$f(p_1 x_1 + p_2 x_2) \leq p_1 f(x_1) + p_2 f(x_2)$$

2.1.2 Ejercicio 2: Esperanza condicional

Utilizando que $\forall f(X) \in \mathcal{H} : E[m(x)f(x)] = E[Yf(x)]$ y $m(x) = E[Y|X = x]$ si $Y - m(x) \perp \mathcal{H}$, con \mathcal{H} el espacio de todas las funciones de X .

Demostrar que vale: $E[\phi(x)Y|X] = \phi(x)E[Y|X]$

2.1.3 Ejercicio 3 (STOP) (SIMULACIÓN)

Sea $X \sim \mathcal{N}(0, 1)$ truncada al intervalo $[-1, 1]$

Imagine $m(x) = E[Y|X = x]$ como:

$$m(x) := \begin{cases} \frac{(x+2)^2}{2} & \text{si } -1 \leq x < -0.5 \\ \frac{x}{2} + 0.875 & \text{si } -0.5 \leq x \leq 0 \\ -5(x - 0.2)^2 + 1.075 & \text{si } 0 < x \leq 0.5 \\ x + 0.125 & \text{si } 0.5 \leq x < 1 \end{cases}$$

Dado un x , la distribución condicional de $Y - m(x)$ es $\mathcal{N}(0, \sigma^2(x))$

Con $\sigma(x) = 0.2 - 0.1 \cos(2x)$

Se pide simular 200 puntos (X, Y) , y graficarlos en un plano.

Además, se piden los 200 pares ordenados en cuestión, para hacer análisis posteriores.

2.1.4 Ejercicio 4 (no entregable)

Dado el problema de decisión introducido en el 'Diseño del receptor de una comunicación binaria', verificar que:

$$\delta(r) = \mathbb{1}\{P(S = 1|R = r) > P(S = 0|R = r)\}$$

maximiza:

$$P(S = \delta(r)) = P(S = \delta(0)|R = 0)P(R = 0) + P(S = \delta(1)|R = 1)P(R = 1)$$

2.2 Clase 2

2.2.1 Ejercicio 5

Terminar la cuenta:

$$P(X, Y) = \int_0^{\frac{1}{2}} (1-x)dx + \int_{\frac{1}{2}}^{\frac{3}{4}} xdx$$

2.2.2 Ejercicio 6

Sean T, F y E exponenciales de intensidad 1 y tenemos que

$$Y = \mathbb{1}\{T + F + E < 7\}$$

Se tenía que E era inobservable. Y se concluyó que:

$$g^*(T, F) = \mathbb{1}\{T + F < 7 - \ln 2\}$$

Se pide calcular el error:

$$L^* = P(g^*(T, F) \neq Y)$$

2.3 Clase 3

2.3.1 Ejercicio 7

Se quería acotar:

$$E[|\tilde{\eta}(x) - \eta(x)|] < \epsilon$$

Para $\epsilon > 0$ existe una función η_ϵ uniformemente continua a valores en el intervalo $[0, 1]$ sobre un conjunto compacto C que se anula fuera de C , y que tiene la siguiente propiedad: $E[|\eta_\epsilon(x) - \eta(x)|] < \epsilon$

Por la desigualdad triangular:

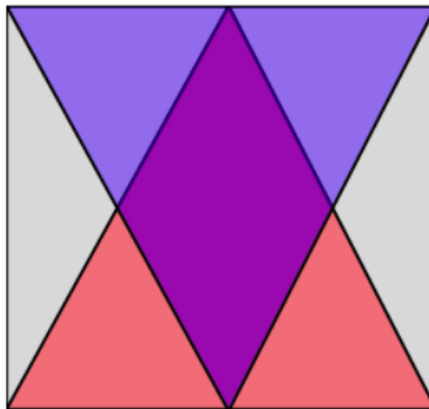
$$E[|\tilde{\eta}(x) - \eta(x)|] \leq E[|\tilde{\eta}(x) - \tilde{\eta}_\epsilon(x)|] + E[|\tilde{\eta}_\epsilon(x) - \eta_\epsilon(x)|] + E[|\eta_\epsilon(x) - \eta(x)|]$$

En clase se acotaron tanto el segundo como el tercer miembro la desigualdad, y el ejercicio es entender por qué el primer término está acotado por el tercero, es decir:

$$E[|\tilde{\eta}(x) - \tilde{\eta}_\epsilon(x)|] \leq E[|\eta_\epsilon(x) - \eta(x)|]$$

2.3.2 Ejercicio 8 (STOP)(SIMULACIÓN)

Dado el siguiente diagrama:



Sean:

X una variable aleatoria con distribución uniforme sobre el triángulo **azul**.

Y una variable aleatoria con distribución uniforme sobre el triángulo **rojo**.

Z se define como: $Z = \frac{1}{2}X + \frac{1}{2}Y$

Se pide:

1. Generar muestras de X y de Y para aprender.
2. Generar muestras a clasificar uniformes sobre el cuadrado completo.
3. Construir una regla del histograma y clasificar las muestras de 2. según lo aprendido en 1. .
4. Graficar los resultados.

2.4 Clase 4

2.4.1 Ejercicio 9: KNN

Se pide simular dos normales bivariadas, ambas con matriz de covarianza $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, pero una centrada en $(-1, 0)$, y la otra en $(1, 0)$.

Se pide:

1. Generar muestras de $N = 100, 1000, \dots$, otros valores. Estas dos serían las clases en las que clasificar los puntos.
2. Generar puntos con una distribución uniforme en $[-4, 4] \times [-4, 4]$.
3. Clasificar los puntos de la uniforme usando KNN , con $K = 1, 3$ y 13 , utilizando las clases del primer ítem.
4. Graficar los resultados.

2.4.2 Ejercicio 9 bis

”Redactar una carilla que contenga la información relevante sobre la distribución normal multivariada. Llevar ejemplos, y mostrar como cambia el gráfico al variar la matriz de covarianza.”

2.4.3 Ejercicio 10

Demostrar que $\forall a, b, c \in \mathbb{R}^+$:

$$(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$$

Sugerencia: usar Cauchy-Schwarz / Jensen.

2.5 Clase 5**2.5.1 Ejercicio 11**

Dadas las desigualdades:

$$P(S_n - E[S_n] \geq \epsilon) \leq e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2}$$

$$P(S_n - E[S_n] \leq -\epsilon) \leq e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2}$$

Probar que se cumple la siguiente desigualdad:

$$P(|S_n - E[S_n]| \leq -\epsilon) \leq 2e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2}$$

Sugerencia: Usar que $P(A \cup B) \leq P(A) + P(B)$

2.5.2 Ejercicio 12

En el ejercicio 11, y a partir de la demostración en clase de la primera de las hipótesis, realizar la demostración de la segunda hipótesis:

$$P(S_n - E[S_n] \leq -\epsilon) \leq e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2}$$

2.5.3 Ejercicio 13

Dada la función:

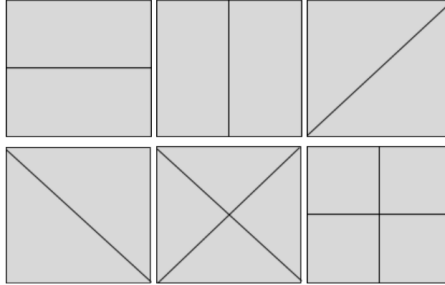
$$\Phi(u) = \ln pe^u + 1 - p - pu$$

Mostrar que:

- $\Phi(0) = 0$
- $\Phi'(0) = 0$

2.5.4 Ejercicio 14 (STOP)

Dados los siguientes clasificadores por particiones:



Elegir cuál es el mejor para las clases y los puntos generados en el ejercicio 8.

2.6 Clase 6**2.6.1 Ejercicio 15**

Demostrar que la siguiente secuencia constituye una serie convergente:

$$a_n = 8(n+1)e^{-n\epsilon^2/32}$$

Y generalizar viendo que $\forall \lambda > 0$:

$$\sum_{n=0}^{\infty} ne^{-\lambda n} < \infty$$

(Sugerencia: Usar que $\int_0^{\infty} xe^{-\lambda x} dx = \frac{1}{\lambda^2}$)

2.6.2 Ejercicio 16

Sea la familia de intervalos:

$$A' = \{(a, b) : a < b\}$$

Luego de calcular $N_{A'}(z_1, z_2) = |\{\emptyset, \{z_1\}, \{z_2\}, \{z_1, z_2\}\}| = 4$ en clase, se pide calcular:

1. $N_{A'}(z_1, z_2, z_3)$
2. $N_{A'}(z_1, z_2, \dots, z_n)$

3 Soluciones

3.1 Clase 1

3.1.1 Ejercicio 1

Solución:

$f(x)$ es convexa en el intervalo $[x_1, x_2]$ si y solo si para cualquier $x_0 = p_1x_1 + p_2x_2$ se cumple la condición a comprobar.

Para comprobar utilizaremos la recta $l(x)$ que contiene a $(x_1, f(x_1))$ y $(x_2, f(x_2)) : \forall x_0 \in [x_1, x_2] : l(x_0) \geq f(x_0)$

Pero $l(x) = mx + b = \frac{f(x_2) - f(x_1)}{x_2 - x_1}x + \left(f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}x_1\right)$

Por simplicidad llamo $y_0 = f(x_1)$ y $y_2 = f(x_2)$

Evaluando:

$$\begin{aligned} l(x_0) &= \frac{y_2 - y_1}{x_2 - x_1}x_0 + \left(y_1 - \frac{y_2 - y_1}{x_2 - x_1}x_1\right) \\ &= \frac{y_2 - y_1}{x_2 - x_1}(p_1x_1 + p_2x_2) + \left(y_1 - \frac{y_2 - y_1}{x_2 - x_1}x_1\right) \end{aligned}$$

Usando que $p_2 = 1 - p_1$

$$\begin{aligned} &= \frac{y_2 - y_1}{x_2 - x_1}(p_1x_1 - p_1x_2) + \frac{y_2 - y_1}{x_2 - x_1}x_2 + \left(y_1 - \frac{y_2 - y_1}{x_2 - x_1}x_1\right) \\ &= \frac{y_2 - y_1}{x_2 - x_1}(x_1 - x_2)(-p_1) + \frac{y_2 - y_1}{x_2 - x_1}(x_2 - x_1) + y_1 \\ &= p_1y_1 - p_1y_2 + y_2 \\ &= p_1y_1 + (1 - p_1)y_2 \\ &= p_1y_1 + p_2y_2 \end{aligned}$$

Finalmente $f(p_1x_1 + p_2x_2) = f(x_0) \leq l(x_0) = p_1f(x_1) + p_2f(x_2)$

3.1.2 Ejercicio 2: Esperanza condicional

Solución:

Sea $h(X) = \frac{f(X)}{g(X)}$, es claro que $h(X)$ puede adoptar cualquier valor de \mathcal{H} , por ejemplo, si busco adoptar $g(X)$ elijo $f(X) = g(X)\phi(X) \in \mathcal{H}$.

Usando la definición de $h(X)$ en la ecuación de hipotesis, se llega a que $\forall h(X) \in \mathcal{H}$:

$$E[m(X)\phi(X)h(X)] = E[Y\phi(X)h(X)]$$

Pero entonces como $m(X)\phi(X) \in \mathcal{H}$, tiene que ser la que minimice la distancia a $Y\phi(X)$, ya que:

$$E[m(X)\phi(X)h(X)] = E[Y\phi(X)h(X)]$$

sii

$$\langle m(X)\phi(X), h(X) \rangle = \langle Y\phi(X), h(X) \rangle$$

sii $\forall h(X) \in \mathcal{H}$:

$$\langle m(X)\phi(X) - Y\phi(X), h(X) \rangle = 0$$

sii

$$m(X)\phi(X) - Y\phi(X) \perp \mathcal{H}$$

.

Por lo tanto $m(X)\phi(X)$ es la esperanza condicional de $Y\phi(X)$ en \mathcal{H}
Finalmente por definición:

$$m(X)\phi(X) = E[Y\phi(X)|X]$$

y con $m(X) = E[Y|X]$ se obtiene la igualdad a demostrar:

$$E[Y|X]\phi(X) = E[Y\phi(X)|X]$$

3.1.3 Ejercicio 3 (STOP) (SIMULACIÓN)

Mirar anexo

3.1.4 Ejercicio 4 (no entregable)

3.2 Clase 2

3.2.1 Ejercicio 5

Solución:

$$\begin{aligned} P(X, Y) &= \int_0^{\frac{1}{2}} (1-x)dx + \int_{\frac{1}{2}}^{\frac{3}{4}} xdx \\ &= \int_0^{\frac{1}{2}} 1dx - \int_0^{\frac{1}{2}} xdx + \int_{\frac{1}{2}}^{\frac{3}{4}} xdx \\ &= \frac{1}{2} - \frac{1}{8} + \left(\frac{9}{32} - \frac{1}{8}\right) \\ &= 17/32 \end{aligned}$$

3.2.2 Ejercicio 6

Solución:

$$\begin{aligned} L^* &= P(g * (T, F) \neq Y) \\ &= P(T + B < 7 - \log 2, T + B + E \geq 7) + P(T + B \geq 7 - \log 2, T + B + E < 7) \\ &= E(e^{-(7-T-B)} \mathbb{1}\{T+B < 7-\log 2\}) + P((1-e^{-(7-T-B)}) \mathbb{1}\{7 > T+B \geq 7-\log 2\}) \\ &= \int_0^{7-\log 2} x e^{-x} e^{-(7-x)} dx + \int_{7-\log 2}^7 x e^{-x} (1 - e^{-(7-x)}) dx \end{aligned}$$

(porque la densidad de $T + B$ es ue^u en $[0, \infty)$)

$$= e^{-7} \left(\frac{(7 - \log 2)^2}{2} + 2(8 - \log 2) - 8 - \frac{7^2}{2} + \frac{(7 - \log 2)^2}{2} \right)$$

(ya que $\int_x^\infty ue^{-u} du = (1+x)e^{-x}$)

$$= 0.0199611\dots$$

3.3 Clase 3

3.3.1 Ejercicio 7

Solución:

Se pide mostrar que:

$$E[|\tilde{\eta}(x) - \tilde{\eta}_\epsilon(x)|] \leq E[|\eta_\epsilon(x) - \eta(x)|]$$

donde:

- $\tilde{f}(x) = E[f(x)|X \in A(x)]$ para toda función f .

Lema previo (desigualdad de Jensen para Esperanzas condicionales):

Sean W, Z variables aleatorias y $f: \mathbb{R} \rightarrow \mathbb{R}$ una función convexa:
 $f(E[W|Z]) \leq E[f(W)|Z]$.

En nuestro ejercicio, sean $Z = X \in A(x)$, $W = \eta_\epsilon(x) - \eta(x)$ y

$$f(x) = |x| := \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}, \text{ claramente convexa en } \mathbb{R}$$

Luego:

$$\tilde{\eta}(x) - \tilde{\eta}_\epsilon(x) = E[\eta(x)|X \in A(x)] - E[\eta_\epsilon(x)|X \in A(x)] = E[\eta(x) - \eta_\epsilon(x)|X \in A(x)]$$

$|\tilde{\eta}(x) - \tilde{\eta}_\epsilon(x)| = f(E[W|Z])$ y además como W y Z son independientes, también lo son $f(W)$ y Z , luego: $E[f(W)|Z] = E[f(W)] = E[|\eta_\epsilon(x) - \eta(x)|]$

Y el ejercicio se reduce a la demostración del **Lema previo**.

3.3.2 Ejercicio 8 (STOP)(SIMULACIÓN)

Mirar anexo

3.4 Clase 4

3.4.1 Ejercicio 9: KNN

Mirar anexo

3.4.2 Ejercicio 9 bis

Normal Multivariada:

Caso general:

Un vector aleatorio $X = [X_1, \dots, X_n]^T$ sigue una distribución normal multivariante si satisface las siguientes condiciones equivalentes:

- Toda combinación lineal $Y = a_1 X_1 + \dots + a_n X_n$ está normalmente distribuida.
- Hay un vector aleatorio $Z = [Z_1, \dots, Z_m]^T$ cuyas componentes son variables aleatorias independientes distribuidas según la normal estándar, un vector $\mu = [\mu_1, \dots, \mu_n]^T$ y una matriz $n \times m$ A tal que $X = AZ + \mu$
- Hay un vector μ y una matriz semidefinida positiva simétrica Σ tal que la función característica de X es:

$$\phi_X(u; \mu, \Sigma) = \exp(i\mu^T u - \frac{1}{2}u^T \Sigma u)$$

Si Σ es una matriz no singular, entonces la distribución puede describirse por la siguiente función de densidad:

$$f_X(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right)$$

donde $|\Sigma|$ es el determinante de Σ . Nótese como la ecuación de arriba se reduce a la distribución normal si Σ es un escalar (es decir, una matriz 1x1).

El vector μ en estas circunstancias es la esperanza de X y la matriz $\Sigma = AA^T$ es la matriz de covarianza de las componentes X_i .

Caso bivalente:

En el caso particular de dos dimensiones, la función de densidad (con media $(0, 0)$) es:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - \frac{2\rho xy}{(\sigma_x\sigma_y)}\right)\right)$$

donde ρ es el coeficiente de correlación entre X e Y . En este caso,

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$$

Interpretación geométrica:

Las curvas de equidensidad de una distribución normal multivariante son elipsoides (es decir, transformaciones lineales de hipersferas) centrados en la media. Las direcciones de los ejes principales de los elipsoides vienen dados por los autovectores de la matriz de covarianza Σ . Las longitudes relativas de los cuadrados de los ejes principales vienen dados por los correspondientes autovectores.

3.4.3 Ejercicio 10**Solución(CS):**

Por CS se ve de forma directa que:

$$3(a^2 + b^2 + c^2) = (1^2 + 1^2 + 1^2)(a^2 + b^2 + c^2) \geq (1^2 a^2 + 1^2 b^2 + 1^2 c^2)^2 = (a + b + c)^2$$

3.5 Clase 5**3.5.1 Ejercicio 11****Solución:**

Por la definición de valor absoluto y la propiedad $P(A \cup B) \leq P(A) + P(B)$ se puede ver directamente el resultado pedido:

$$\begin{aligned} P(|S_n - E[S_n]| \leq -\epsilon) &= P((S_n - E[S_n] \geq \epsilon) \cup (S_n - E[S_n] \leq -\epsilon)) \\ &\leq P(S_n - E[S_n] \geq \epsilon) + P(S_n - E[S_n] \leq -\epsilon) \\ &= e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2} + e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2} \\ &= 2e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2} \end{aligned}$$

3.5.2 Ejercicio 12**Solución:**

Sabemos que, si X_1, X_2, \dots, X_n son variables aleatorias acotadas e independientes tales que $P(a_i \leq X_i \leq b_i) = 1$, con $i = 1, 2, \dots, n$, se cumple:

$$P(S_n - E[S_n] \geq \epsilon) \leq e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2} \quad (hip)$$

Sea $X_i^* = -X_i$ para todo $i = 1, \dots, n$, luego $P(-b_i \leq X_i^* \leq -a_i) = 1$

Además defino $S_n^* = \sum_{i=1}^n X_i^*$

Como las variables $X_1^*, X_2^*, \dots, X_n^*$ cumplen con las condiciones de la hipótesis:

$$P(S_n^* - E[S_n^*] \geq \epsilon) \stackrel{(hip)}{\leq} e^{-2\epsilon^2 / \sum_{i=1}^n ((-a_i) - (-b_i))^2}$$

Pero:

- $S_n^* = \sum_{i=1}^n X_i^* = \sum_{i=1}^n (-X_i) = -S_n$
- $E[S_n^*] = E[\sum_{i=1}^n X_i^*] = E[\sum_{i=1}^n (-X_i)] = -E[-S_n] = -E[S_n]$
- $((-a_i) - (-b_i))^2 = (b_i - a_i)^2$

Luego:

$$P(S_n - E[S_n] \leq -\epsilon) = P(S_n^* - E[S_n^*] \geq \epsilon) \leq e^{-2\epsilon^2 / \sum_{i=1}^n ((-a_i) - (-b_i))^2} = e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2}$$

$$P(S_n - E[S_n] \leq -\epsilon) \leq e^{-2\epsilon^2 / \sum_{i=1}^n (b_i - a_i)^2}$$

3.5.3 Ejercicio 13**Solución:**

$$\Phi(0) = \ln(1) = 0$$

$$\Phi'(u) = \frac{(pe^u)}{(pe^u - p + 1)} - p, \quad \Phi'(0) = \frac{(pe^0)}{(pe^0 - p + 1)} - p = p - p = 0$$

3.5.4 Ejercicio 14 (STOP)

Mirar anexo

3.6 Clase 6**3.6.1 Ejercicio 15****Solución:**

Teorema: Si f es positiva, continua y decreciente para $x \geq 1$ y $a_n = f(n)$, entonces

$$\sum_{n=0}^{\infty} a_n \quad \text{y} \quad \int_0^{\infty} f(x) dx$$

convergen o divergen ambas simultáneamente.

Si $T \sim \exp(\lambda)$ y $a_n = ne^{-\lambda n}$,
es inmediato que como $\int_0^{\infty} xe^{-\lambda x} dx = \frac{1}{\lambda} E[T] = \frac{1}{\lambda^2} < \infty$
entonces la serie asociada es convergente.

3.6.2 Ejercicio 16**Solución:**

El problema equivale a contar la cantidad de subconjuntos con elementos consecutivos hay en un conjunto $\{z_1, z_2, \dots, z_n\}$ con $z_1 < z_2 < \dots < z_n$.

Los subconjuntos con mas de un elemento quedan completamente definidos simplemente eligiendo sus elementos extremos,

Luego hay $\binom{n}{2} = \frac{n(n-1)}{2}$ de estos, sumados a los n conjuntos de un elemento y al conjunto vacío, se tiene que:

$$N_{A'}(z_1, z_2, \dots, z_n) = \frac{n(n-1)}{2} + n + 1$$

En caso $n = 3$ es inmediato habiendo deducido este, por lo tanto.

$$N_{A'}(z_1, z_2, z_3) = \frac{3(3-1)}{2} + 3 + 1 = 7.$$

Que coincide con el cardinal del conjunto (hallado de forma exhaustiva):

$$\{\emptyset, \{z_1\}, \{z_2\}, \{z_3\}, \{z_1, z_2\}, \{z_2, z_3\}, \{z_1, z_2, z_3\}\}$$