Inferencia II 2016
*Práctico 1*

# Ejercicio 1

Sea $X$ una variable aleatoria con distribución Gamma$(\alpha, \beta)$.

1. Hallar la distribución de $Y = 1/X$. A esta distribución se le llama *Gamma-Inversa*$(\alpha, \beta)$ y se denota $Y \sim$ Inv-Gamma$(\alpha, \beta)$.
2. Calcular E$(Y)$ y Var$(Y)$ (especificar los valores de $\alpha$ y $\beta$ para los cuales existen estos momentos).
3. Simular[1] $n = 1000$ observaciones de $Y$ y calcular $\widehat{E}(Y)$ y $\widehat{Var}(Y)$ (elegir los valores de $\alpha$ y $\beta$).
4. Graficar conjuntamente las densidades de $X$ e $Y$ para los siguientes casos: $(\alpha, \beta) = (0.1, 0.1), (2, 5), (5, 0.5)$ y $(10, 10)$.

# Ejercicio 2

1. Dada la variable aleatoria $X$ con fgm $M_X(t) = E(e^{tX})$ encontrar la fgm de $\sum_{i=1}^{n} X_i$ donde $X_1, \ldots, X_n$ es una muestra *i.i.d* de $X$.
2. ¿Si $X \sim Exp(\lambda)$ cuál es la distribución de $\sum_{i=1}^{n} X_i$?

# Ejercicio 3

Mostrar que para dos variables aleatorias $X$ e $Y$ se cumple:

1. $E(X) = E(E(X|Y))$
2. $Var(X) = E(Var(X|Y)) + Var(E(X|Y))$

# Ejercicio 4

Suponiendo que $Y|\theta \sim$Binomial$(n, \theta)$ y que $\theta \sim$ Beta$(\alpha, \beta)$.

1. Hallar $E(Y)$ y $Var(Y)$.
2. Encontrar la distribución marginal de $Y$ (conocida como distribución *Beta-Binomial).* Observar en particular el caso $\alpha = \beta = 1$.

[1] En $R$ puede utilizar la función *rgamma()*

3. Obtener $N = 1000$ valores simulados de $Y$ utilizando las distribuciones de $\theta$ y de $Y|\theta$ para $n = 25$, $\alpha = 5$ y $\beta = 15$. Calcular $\widehat{E}(Y)$ y $\widehat{Var}(Y)$ y hacer un gráfico de frecuencias de $y_1, \ldots, y_n$.

4. Obtener la distribución de $\theta|Y = y$ y deducir $E(\theta|Y)$ $\text{Var}(\theta|Y)$.

# Ejercicio 5

Suponiendo que $Y|\lambda \sim$Poisson($\lambda$) y que $\lambda \sim$ Gamma$(\alpha, \beta)$.

1. Hallar la distribución de $Y$, $E(Y)$ y $Var(Y)$.

2. Si $\alpha \in \mathbb{N}$ reconocer la distribución de $Y$.

3. Hallar la distribución de $\lambda|Y$.

4. Suponiendo ahora que $Y|N \sim$ Binomial$(N, p)$, $N|\lambda \sim$ Poisson$(\lambda)$ y que $\lambda \sim$ Gamma$(\alpha, \beta)$. Mostrar que la distribución de $Y$ es la misma que la hallada en 1 cuando $p = 1$.

# Ejercicio 6

Una generalización de la distribución Beta es la distribución de *Dirichlet*. La versión bivariada de esta distribución tiene la siguiente función de densidad:

$$p(x,y) = \begin{cases} Cx^{a-1}y^{b-1}(1-x-y)^{c-1} & \text{si } 0 < x < 1 \quad 0 < y < 1 \quad 0 < y < 1 - x < 1 \\ \\ 0 & \text{en otro caso} \end{cases}$$

donde $a > 0$, $b > 0$ y $c > 0$ son constantes.

1. Mostrar que $C = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}$.

2. Mostrar que, marginalmente, ambas $X$ y $Y$ tienen una distribución Beta.

3. Encontrar la distribución de $Y|X = x$.

4. Calcular $\text{Cov}(X, Y)$ y $\rho(X, Y)$.

# Ejercicio 7

Una generalización del concepto de variables aleatorias independientes e idénticamente distribuidas es el de variables aleatorias *intercambiables* desarrollado por De Finetti:

*Las variables aleatorias $Y_1, \ldots, Y_n$ son intercambiables si cualquier permutación de cualquier subconjunto de ellas de tamaño $k$ ($k \leq n$) tiene exactamente la misma distribución.*

Sean $Y_1, \ldots, Y_n$ variables aleatorias tales que **condicionalmente** en $p$ son independientes e idénticamente distribuidas con $Y_i|p \sim$ Bernoulli$(p)$, $i = 1, \ldots, n$. Asuma que $p \sim$ Uniforme(0,1).

1. Muestre que la distribución marginal de cualquier subconjunto de tamaño $k$, $(0 < k \leq n)$, de $Y_1, \ldots, Y_n$ es la misma que

$$P(Y_1 = y_1, \ldots, Y_k = y_k) = \frac{t!(k-t)!}{(k+1)!},$$

donde $t = \sum_{i=1}^{k} y_i$. Lo cual muestra que $Y_1, \ldots, Y_n$ son intercambiables.

2. Muestre que, marginalmente,

$$P(Y_1 = y_1, \ldots, Y_n = y_n) \neq \prod_{i=1}^{n} P(Y_i = y_i),$$

lo cual implica que $Y_1, \ldots, Y_n$ son intercambiables pero no independientes.

# Ejercicio 8

Suponga que se observa $y \sim$Binomial($n$,$p$). El intervalo de confianza estándar para $p$ al 90 % está dado por

$$C(y) = \left( \hat{p} - 1{,}645\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 1{,}645\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right),$$

donde $\hat{p} = y/n$.

Este intervalo se construye bajo el supuesto de que

$$P\big(p \in C(y)\big) = 0{,}90 \quad \forall p \in (0,1).$$

1. Escriba una función en R que calcule los límites de un intervalo de confianza para $p$ al 90 %. Los parámetros de la función son $y$ y $n$.
2. Suponga que $n = 20$ y que el verdadero valor de $p$ es 0,5. Para este caso en particular, por medio de simulación, estudie cual es la verdadera probabilidad de cobertura de $C(y)$.
3. Repita el punto anterior, pero ahora suponiendo que el verdadero valor de $p$ es 0,05.
4. Repita los dos puntos anteriores pero ahora suponiendo que $n = 100$. ¿Cuáles son sus conclusiones?

# Ejercicio 9

Suponga que $y|\theta \sim$Gamma$(\alpha, \theta)$, para $\alpha$ conocido.

¿Existe una priori conjugada en este caso? En caso afirmativo, identifíquela y encuentre la posteriori.

# Ejercicio 10

Suponga que $y|\theta \sim$Gamma$(\theta, \beta)$, para $\beta$ conocido.

¿Existe una priori conjugada en este caso? En caso afirmativo, identifíquela y encuentre la posteriori.

# Ejercicio 11

Suponga que la prevalencia de una enfermedad es $p$, es decir, si elijo una persona al azar: $p(enfermo) = p$ y $p(sano) = 1 - p$. Puede pensarse que el parámetro $\theta$ toma los valores $enfermo$ y $sano$ con esa priori.

Un análisis para detectar la enfermedad indica con $y = 0$ que "la persona está sana" y con $y = 1$ que "la persona está enferma", pero este análisis puede equivocarse. La probabilidad de falso positivo es $p(y = 1|sano) = 0{,}04$ y de falso negativo es $p(y = 0|enfermo) = 0{,}02$.

1. Una persona se hace el análisis por primera vez y le da $y_1 = 0$. Calcule la posteriori de $\theta$.
2. La persona se hace el análisis por segunda vez, y le da $y_2 = 0$. Calcule la nueva posteriori de dos maneras:

   *a*) Usando la información de los dos análisis conjuntamente.

   *b*) Usando la posteriori calculada en 1) como priori para la segunda observación.

3. Calcule las posterioris teniendo una y dos observaciones para $p = 0{,}10$.