III) Sea $0 < P(A) < 1$, y sea I_A la función indicatriz de A, es decir, $I_A(x) = 1$ si $x \in A$, $I_A(x) = 0$ si $x \in A^c$. Recordando que $E(I_A) = P(A)$, calcular:

1) El coeficiente de correlación $\rho(I_A, I_A)$ (4 puntos).

2) $\rho(I_A, I_{A^c})$ (4 puntos).

3) $\rho(I_A, I_B)$, donde A y B son eventos independientes (2 puntos).

$$1)\quad \rho(I_A, I_A) = \frac{E I_A I_A - E I_A E I_A}{\sqrt{Var I_A}\sqrt{Var I_A}}$$

$$= \frac{PA - (PA)^2}{PA - (PA)^2} = 1$$

$$2)\quad \rho(I_A, I_{A^c}) = \frac{E I_A I_{A^c} - E I_A E I_{A^c}}{\sqrt{PA(1-PA)}\sqrt{P(A^c)(1-P(A^c))}}$$

$$= -\frac{PA\, PA^c}{PA\, PA^c} = -1$$

$$3)\quad \rho(I_A, I_B) = \frac{E I_A I_B - E I_A E I_B}{\sqrt{Var I_A}\sqrt{Var I_B}} = 0$$

IV) En una ciudad donde diáriamente circulan 10 000 automóviles, la probabilidad de que cada coche sufra un accidente en un día dado es 1/5000. Suponiendo que los accidentes ocurren de manera independiente, aproximar la probabilidad de que como máximo se produzca un accidente en el día escogido, de las siguientes formas:

1) Usando la distribución de Poisson (4 puntos).
2) Usando la distribución normal (4 puntos).
3) Decidir razonadamente que aproximación es mejor (2 puntos).

1) $\lambda = np = 2$, $P(Y=0)+P(Y=1) = e^{-2} + e^{-2}\frac{2}{1!} \approx 0'406006$
$Y \sim \text{Poisson}(2)$

2) Sea $S_n = \sum_{i=1}^{n} X_i$, $P(X_i = 1) = \frac{1}{5000}$. Aproximamos

$P(S_{10000} \leq 1) = P(S_{10000} \leq 1'5)$ (corrección de de Moivre-Laplace)

$= P\left(\frac{S_{10000} - 2}{\sqrt{2 - \frac{1}{2500}}} \leq -\frac{1}{2\sqrt{2-\frac{1}{2500}}}\right)$ usando

$P(Z \leq -0'3536) = 1 - P(Z \geq -0'3536)$

$= 1 - P(Z \leq 0'3536) \approx 1 - 0'6368 = 0'3632$

Sin la corrección, tenemos

$P(Z \leq -0'7072) = 1 - P(Z \leq 0'7072) \approx 1 - 0'7611$

$= 0'2389$

3) La aproximación de Poisson es muy buena, al tener $Y \sim \text{Poisson}(2)$ la misma media que S_{10000} (por elección de λ) y varianza muy similar: 2 frente a $2 - \frac{1}{2500}$.

La aproximación normal es mala. Al ser $p \approx 0$, la convergencia hacia la normal requiere un valor de n mucho más grande que cuando $p \approx \frac{1}{2}$.