- **No borre, tache pero muestre su trabajo.**
- **Justifique sus respuestas**.
- **Duración de examen: 135 minutos**.
- Total: 104/104 puntos.

Resultados

- X variable aleatoria que sigue la *distribución geométrica* con parámetro $0 < p < 1$, entonces $X \sim \text{Geom}(p)$ tiene función masa de probabilidad f y función generadora de momentos m dadas por

$$f(x) = p(1-p)^x \mathbb{1}_{\{0,1,\dots\}}(x); \quad m(t) = \frac{p}{1-(1-p)e^t},\ t \leq -\log(1-p)$$

- X variable aleatoria que sigue la *distribución Poisson* con parámetro $\lambda > 0$, entonces $X \sim \text{Po}(\lambda)$ tiene función masa de probabilidad f y función generadora de momentos m dadas por

$$f(x) = \lambda^x e^{-\lambda}/x! \mathbb{1}_{\{0,1,\dots\}}(x); \quad m(t) = e^{\lambda(e^t-1)},\ t \in \mathbb{R}$$

- Y variable aleatoria que sigue la *distribución Gamma* con parámetro de forma $\alpha > 0$ y parámetro tasa $\lambda > 0$, tiene función de densidad

$$f(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} \mathbb{1}_{(0,\infty)}(y)$$

  Luego, su función generadora de momentos es $m(t) = (1 - t/\lambda)^{-\alpha}$ para $t < \lambda$ y su r–ésimo momento

$$\mathbb{E}[Y^r] = \frac{\Gamma(\alpha+r)}{\lambda^r \Gamma(\alpha)}, \quad r > -\alpha$$

- Las distribuciones *Exponencial* y *Ji-cuadrada (χ^2)* son casos particulares de la distribución Gamma.

- Sean X_1 y X_2 subvectores del vector aleatorio $X \sim \text{N}_n(\mu, \Sigma)$. Esto es, si

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \text{N}_n\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

  donde μ_i y Σ_{ij} son la correspondientes particiones del vector de medias μ y matriz de covarianzas Σ. Se tiene entonces que,

$$(X_1 | X_2 = x_2) \sim \text{N}_{n_1}(\mu_{1|2}, \Sigma_{1|2})$$

  donde el vector de medias $\mu_{1|2}$ y la matriz de covarianzas $\Sigma_{1|2}$ están dados por

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) \quad \text{y} \quad \Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

- *Teorema Central del Límite*: Sea $X_1, X_2, \dots$ una sucesión de sucesión de variables aleatorias independientes e idénticamente distribuidas (*v.a.i.i.d.*) con media μ y varianza σ^2, $S_n = X_1 + \cdots + X_n$ y $\bar{X}_n = \frac{1}{n} S_n$, entonces para $Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} = \dfrac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$,

$$Z_n \xrightarrow{D} Z \sim \text{N}(0,1)$$

- Sean $Z \sim \text{N}(0,1)$ y $\Phi(z) = P(Z \leq z)$. Se tiene entonces que

$$\Phi(1/3) = 0.631, \quad \Phi(1/2) = 0.691, \quad \Phi(1) = 0.841, \quad \Phi(2) = 0.977, \quad \Phi(3) = 0.999$$

$$\Phi(1.282) = 0.900, \quad \Phi(1.645) = 0.950, \quad \Phi(1.960) = 0.975 \quad \Phi(2.326) = 0.990$$

Responda cinco problemas

1. [**20 puntos**] Se tiene una colección de n problemas de los cuales m de ellos se usaron en el examen final del periodo de mayo. Para el periodo de diciembre se seleccionaron al azar m problemas y para el periodo de enero nuevamente se han elegido al azar m problemas. Sea D la variable aleatoria que denota el número de problemas en diciembre que no fueron usados en el periodo de mayo y E que denota el número de problemas del periodo de enero no usados anteriormente. Suponga que $n \geq 3m$.

   Responda los siguientes incisos. Su respuesta debe quedar expresada en términos de los parámetros n y m.

   a) [4 puntos] Determine f_D, la función masa de probabilidad marginal de D.

   b) [10 puntos] Determine f_E, la función masa de probabilidad marginal de E.

   c) [6 puntos] Si sabe que e de los problemas seleccionados para enero no fueron utilizados anteriormente, determine la probabilidad de que d de los m problemas de diciembre hayan sido usados en el periodo de mayo.

2. [**20 puntos**] Sea (X, Y) un vector aleatorio que sigue una distribución normal bivariada tal que marginalmente X y Y siguen la distribución normal estándar y tienen una correlación ρ. Defina ahora $U = Y - X$ y $V = Y + X$,

   a) [16 puntos] Determine la distribución conjunta del vector (U, V).

   b) [4 puntos] Indique si las componentes U y V son independientes. Justifique sus respuestas.

3. [**20 puntos**] Sea $X_1, X_2, \ldots,$ una sucesión de *v.a.i.i.d.*'s con función generadora de momentos M_X y N una variable entera positiva, independiente de los X_i's, con función masa de probabilidad f_N. Determine la *f. g. m.* $M_Y(t)$, donde

$$Y = \sum_{i=1}^{N} X_i$$

   Si las X_i siguen la distribución exponencial de media $\theta = 2$ y N tiene la *f. m. p.* dada por la tabla

n	1	2	3
p	1/6	3/6	2/6

   Evalúe $M_Y(1/4)$.

4. [**20 puntos**]

   a) [10 puntos] Considere X, Y y Z, *v.a.i.i.d.*'s continuas. Muestre que

$$\mathbb{P}(X < \text{mín}\{Y, Z\}) + \mathbb{P}(Y < \text{mín}\{X, Z\}) + \mathbb{P}(Z < \text{mín}\{X, Y\}) = 1$$

   b) [4 puntos] Sean X y Y *v.a.i.i.d.*'s que siguen la distribución Bernoulli con probabilidad de éxito p. Esto es, $X, Y \sim \text{Be}(p)$. Encuentre la función masa de probabilidad conjunta del vector aleatorio (U, V), donde $U = \text{mín}\{X, Y\}$ y $V = \text{máx}\{X, Y\}$. Calcule $\mathbb{E}[U]$ y $\mathbb{E}[V]$.

5. [**20 puntos**] En cierta zona de la ciudad el transporte público lo cubren dos compañías, *Seg* y *Lux*. La compañía S se caracteriza porque sus autobuses pasan por el parador cada 10 minutos. Por otro lado, L ofrece autobuses más cómodos pero el tiempo de arribo al parador es aleatorio distribuido exponencialmente con una media de 10 minutos. Un usuario, Alfredo, llega al parador de manera aleatoria.

   Responda los siguientes incisos expresando los resultados numéricos con al menos 3 decimales.

   a) [10 puntos] Determine la probabilidad de que el autobús de *Lux* arribe antes que el de *Seg*.

   b) [10 puntos] Determine la función de distribución del tiempo de espera de Alfredo.

6. [**20 puntos**]

   a) [8 puntos] Sean X y Y variables aleatorias y a, b, c, d constantes reales. Si se definen $U = a + bX$ y $V = c + dY$, muestre que $\operatorname{corr}(U, V) = \operatorname{corr}(X, Y)$.

   b) [12 puntos] Muestre que para cualquier función h, la variable aleatoria $(Y - \mathbb{E}[Y|X])$ no está correlacionada con $h(X)$. $\big($*Sugerencia*: note que $h(X)\mathbb{E}[Y|X] = \mathbb{E}[h(X)Y|X]$.$\big)$

   La variable $(Y - \mathbb{E}[Y|X])$ es llamada el *residual* de usar X para *predecir* Y,

7. [**20 puntos**] Sean $X_1, \ldots, X_n$, *v.a.i.i.d.*'s con distribución común uniforme en el intervalo $(0, 1)$. Demuestre que

$$2 \log \left(\frac{1}{\prod_{i=1}^{n} X_i} \right) \sim \chi^2_{2n}$$

   Sugerencia: ¿Cómo se distribuye $-2 \log X_i$?

8. [**20 puntos**]

   a) [10 puntos] Una compañía de golosinas afirma que su producto tiene un contenido medio de 2.0 gr. y una desviación estándar de 1.2 gr. de azúcar por porción. Se ha tomado una muestra aleatoria de 100 porciones y se determinó que el promedio del contenido de azúcar es de 2.4 gr/porción. ¿Respaldaría usted lo afirmado por la compañía? Justifique su respuesta.

   b) [10 puntos] Si se deseara *estimar* (aproximar) con el promedio el valor real del contenido medio de azúcar por porción, con un error no mayor a 0.3 gr. y una *confianza* del 95 % o mayor, ¿de qué tamaño debería ser el tamaño de la muestra? Esto es, determine n para que el error de estimación se cometa no más de 5 % de las ocasiones si se repitiese muchas veces la estimación de la media.