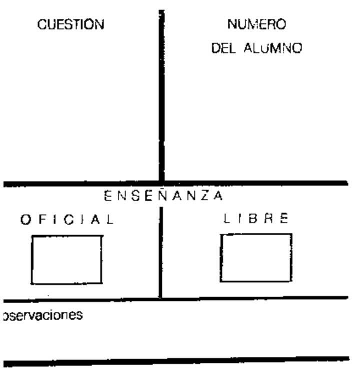
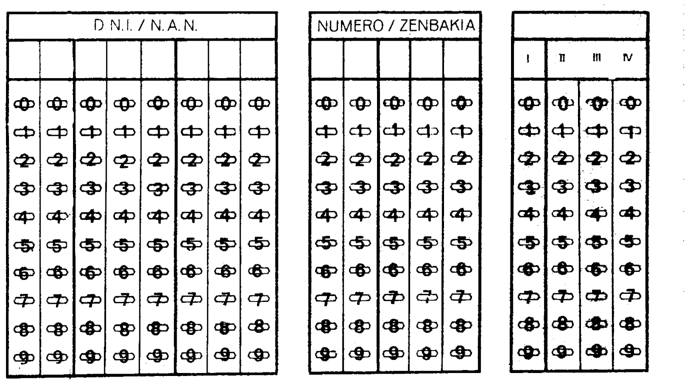
INSTRUCCIONES

1. El examen consta de cuestiones, que se responden sobre la hoja de codificación proporcionada, y problemas, que se responden en papel aparte.
2. Para escoger una respuesta, basta efectuar una marca **rellenando debidamente el rectángulo sobre el que está la letra escogida** en la hoja de codificación. Piénsalo antes; aunque puedes borrar si escribes con lápiz (número 2 o similar), marcas que no estén perfectamente borradas pueden ser leídas. Te aconsejamos que señales sobre el formulario de examen las respuestas que te parezcan adecuadas, y emplees los últimos diez minutos del tiempo asignado en transcribirlas a la hoja de codificación.
3. Hay siempre, en las preguntas de elección múltiple, una **única** respuesta correcta. Todas las cuestiones correctamente resueltas valen 1 punto mientras que las fallidas suponen una penalización de 0.2 puntos. Las preguntas no contestadas no suponen penalización.
4. Cada uno de los problemas, A, B y C debe responderse en una hoja de papel diferente. La recogida se producirá escalonadamente, en los momentos que constarán en la pizarra; primero, la hoja de codificación, y luego los problemas A, B, y C en este orden.
5. El formulario de examen tiene seis hojas numeradas correlativamente al pie (del 0.1 al 0.6). Cerciórate de recibirlas todas, y reclama si tu formulario fuera incompleto. Hay distintos tipos de examen. Este es del tipo 0; marca un 0 en la columna I de tu hoja de codificación, como en el ejemplo.
6. Los puntos obtenibles en cuestiones y problemas son 30 y 30 respectivamente. Son precisos 15 y 15 para superar el examen. Cuestionarios con puntuaciones iguales o superiores a 14 pueden en algún caso ser compensados por una buena nota en los problemas.
7. Rellena tus datos en la hoja de codificación y pliegos de papel suministrados. En "Convocatorias" (columna II) pondrás el número de convocatorias consumidas *incluyendo ésta.*

Ejemplo:

$\underbrace{12545}$ PEREZ, Ernesto Examen tipo $\underbrace{0}$ Convocatorias





CUESTIONES (Duración: 1 hora 40 minutos)

1. **PREGUNTA–REGALO.** La capital de España es:

   (A) París (B) Sebastopol (C) Madrid (D) Londres (E) Pekín

Las cuestiones 2 a 4 hacen referencia al siguiente enunciado:

El área en metros cuadrados de los pisos que se encuentran a la venta en una zona residencial sigue una distribución normal $N(90, \sigma = 10)$. Se consideran pisos grandes aquellos cuya área sea de al menos 98.4 metros cuadrados. Se asume independencia entre las áreas de los distintos pisos. **Nota**: Al calcular esta probabilidad debes redondear a un decimal.

2. Si se toma una muestra aleatoria de 10 pisos en dicha zona, la probabilidad de que al menos 6 de ellos sean grandes es:

   (A) 0.9991 (B) 0.9936 (C) 0.0064 (D) 0.0055 (E) 0.0009

3. En la misma muestra de 10 pisos, la probabilidad de que exactamente 8 pisos **no** sean grandes es:

   (A) 1 (B) 0.2013 (C) 0.0001 (D) 0.3020 (E) 0.6778

4. Si ahora se toma una muestra aleatoria de 150 pisos en dicha zona, la probabilidad aproximada de que al menos 35 de ellos sean grandes es:

   (A) 0.2061 (B) 0.4286 (C) 0.8212 (D) 0.5714 (E) 0.1788

Las cuestiones 5 a 7 hacen referencia al siguiente enunciado:

El número de alumnos que acude por minuto a la secretaría de una determinada Universidad sigue una distribución de Poisson con media igual a 1. Se asume independencia entre las llegadas de alumnos en diferentes minutos.

5. La probabilidad de que en un minuto acudan a la secretaría menos de 2 alumnos es:

   (A) 0.3679 (B) 0.0803 (C) 0.1839 (D) 0.9197 (E) 0.7358

6. La probabilidad de que en el transcurso de 5 minutos acudan a la secretaría al menos 6 alumnos es:

   (A) 0.7622 (B) 0.2378 (C) 0.1334 (D) 0.6160 (E) 0.3840

7. La probabilidad aproximada de que en una hora acudan a la secretaría exactamente 60 alumnos es:

   (A) 0.0956 (B) 0.0478 (C) 0.1478 (D) 0.5239 (E) 0.0239

8. Sea $\{X_n\}_{n \in \mathcal{N}}$ una sucesión de variables aleatorias con distribución $N(1, \sigma^2 = 1/n^3)$. La sucesión converge:

   (A) En distribución, probabilidad y media cuadrática a $X = 1$

   (B) Sólo en probabilidad a $X = 1$

   (C) Sólo en distribución a $X = 1$

   (D) Sólo en distribución y probabilidad a $X = 1$

   (E) Todo falso

9. Sea $\{X_n\}_{n\geq 2}$ una sucesión de variables aleatorias definidas como:

$$P(X_n=-n^2)=\frac{1}{n^3};\qquad P(X_n=0)=1-\frac{2}{n^3};\qquad P(X_n=n^2)=\frac{1}{n^3}$$

La sucesión converge:

(A) Sólo en probabilidad a $X=0$

(B) Sólo en distribución a $X=0$

(C) Sólo en distribución y probabilidad a $X=0$

(D) En distribución, probabilidad y media cuadrática a $X=0$

(E) En distribución y media cuadrática a $X=0$

10. Sean X, Y y Z v.a. independientes tales que $X\in\gamma(1/2,1)$, $Y\in\gamma(1/2,2)$ y $Z\in\gamma(2,3)$. La distribución de la v.a. $X+Y+Z$ es:

(A) $\chi^2_{\overline{6}|}$ (B) $\gamma(1/2,3)$ (C) $\gamma(2,6)$ (D) Todo falso (E) $\gamma(1/2,6)$

11. Sean X e Y v.a. independientes, ambas con distribución $N(0,1)$. Sea $Z=(X/Y)^2$. La distribución de la v.a. Z es:

(A) t_1 (B) $N(0,1)$ (C) $\chi^2_{\overline{1}|}$ (D) $\mathcal{F}_{\overline{1,1}|}$ (E) $Z\equiv 1$

12. Sea X una variable aleatoria con distribución $t_{\overline{25}|}$. Entonces $P(-2.48<X<-0.684)$ es:

(A) 0.02 (B) 0.50 (C) 0.24 (D) 0.56 (E) 0.74

13. Sea X una variable aleatoria con distribución exponencial de media $\frac{1}{2}$. Entonces, el valor aproximado de $P(-1.5<X<1/4)$ es:

(A) 0.39 (B) 0.86 (C) 0.14 (D) 0.22 (E) 0.61

14. Sea X una variable aleatoria con distribución $\mathcal{F}_{5,8}$. Entonces el valor de k tal que $P(X>k)=0.95$ es:

(A) 0.21 (B) 3.69 (C) 0.27 (D) 3.34 (E) 4.82

Las cuestiones 15 y 16 hacen referencia al siguiente enunciado:

En un juego de tiro al blanco un jugador lanza dardos de forma continua e idéntica sobre el objetivo. Se desea estimar la probabilidad θ de que un dardo dé en el blanco.

15. Si el jugador ha necesitado 13 lanzamientos para acertar sobre el objetivo, la estimación de θ por el método de máxima verosimilitud es:

(A) $\frac{12}{13}$ (B) $\frac{1}{13}$ (C) $\frac{1}{12}$ (D) $\frac{13}{14}$ (E) $\frac{11}{12}$

16. Después del objetivo que requirió 13 lanzamientos para acertar, el jugador lo vuelve a intentar y, en esta segunda ocasión requiere 15 lanzamientos para acertar sobre este segundo objetivo. La nueva estimación de θ por el método de máxima verosimilitud considerando todos los lanzamientos es:

(A) $\frac{14}{15}$ (B) $\frac{1}{14}$ (C) $\frac{1}{28}$ (D) $\frac{1}{17}$ (E) $\frac{12}{13}$

Las cuestiones 17 a 20 hacen referencia al siguiente enunciado:

Sea X una variable aleatoria con distribución uniforme $U[\theta-1,\theta+1]$ de la que, para estimar el parámetro θ, se toma una m.a.s. de tamaño n, $X_1,\ldots,X_n$. Se propone el estimador $\hat{\theta}=\overline{X}+\frac{1}{n^2}$. Además, se sabe que la varianza de una v.a. $U[a,b]$ es $\sigma^2=(b-a)^2/12$.

17. ¿Es este estimador insesgado?

(A) No (B) - (C) Sí (D) - (E) -

18. La varianza de este estimador es:

(A) $\frac{1}{3}$ (B) $\frac{1}{12n}$ (C) $\frac{1}{3n}$ (D) $\frac{1}{12n}+\frac{1}{n^2}$ (E) $\frac{1}{12}$

19. El error cuadrático medio de este estimador es:

(A) $\frac{1}{12n}$ (B) $\frac{1}{3n}+\frac{1}{n^4}$ (C) $\frac{1}{12}$ (D) $\frac{1}{12n}+\frac{1}{n^2}$ (E) $\frac{1}{3}+\frac{1}{n^4}$

20. ¿Es este estimador consistente?

(A) Sí (B) - (C) - (D) - (E) No

Las cuestiones 21 a 23 hacen referencia al siguiente enunciado:

Sea X una v.a. con función de densidad

$$f(x;\theta)=\theta^2x^{\theta^2-1},\qquad 0<x\leq 1,\qquad \theta>0$$

Se quiere contrastar la hipótesis nula $H_0:\theta=1$ frente a la hipótesis alternativa $H_1:\theta=2$. Para ello se ha tomado de dicha población una muestra de un sólo elemento (es decir, se observa X).

21. La región crítica de máxima potencia para X es de la forma:

(A) $(0,K)$ (B) (K_1,K_2) (C) $(K,1)$ (D) $(K_1,K_2)^{\mathrm{c}}$ (E) Todo falso

22. Si $\alpha=0.10$, la región crítica de mayor potencia es:

(A) $(0.95,1)$ (B) $(0.10,0.90)$ (C) $(0.90,1)$ (D) $[0.10,0.90]^c$ (E) $(0,0.10)$

23. La potencia del contraste es:

(A) 0.0975 (B) 0.3439 (C) 0.9025 (D) 0.6561 (E) 0.0025

Las cuestiones 24 y 25 hacen referencia al siguiente enunciado:

Sea $X_1,\ldots,X_n$ una m.a.s. de tamaño 20 tomada de una población binaria $b(p)$. Queremos contrastar la hipótesis nula $H_0:p=0.20$ frente a la hipótesis alternativa $H_1:p=0.40$. Sea $Z=\sum_{i=1}^n X_i$, el número total de aciertos en las 20 repeticiones del experimento binario.

24. Al nivel de significación del 5%, la región crítica de mayor potencia es:

(A) $Z\geq 8$ (B) $Z\in[2,8]$ (C) $Z\geq 7$ (D) $Z\in[2,8]^{\mathrm{c}}$ (E) $Z\leq 2$

25. La potencia del contraste es:

(A) 0.4159 (B) 0.5956 (C) 0.5841 (D) 0.4044 (E) 0.2447

Las cuestiones 26 y 27 hacen referencia al siguiente enunciado:

En un proceso de fabricación se desea contrastar la hipótesis nula de que el porcentaje de piezas no defectuosas fabricadas es al menos de un 10%, frente a la alternativa de que es menor de un 10%. Se toma una m.a.s. de tamaño 300, en donde se observan 24 piezas no defectuosas.

26. Con una confianza aproximada del 95% se puede afirmar que el porcentaje de piezas no defectuosas fabricadas se encuentra en el intervalo:

(A) $(0.02,\ 0.14)$ (B) $(0,\ 0.15)$ (C) $(0.01,\ 0.15)$ (D) $(0,\ 0.08)$ (E) $(0.05,\ 0.11)$

27. Al nivel de significación aproximado del 5%, la decisión del contraste es:

(A) No rechazar la hipótesis nula (B) - (C) Rechazar la hipótesis nula (D) - (E) -

Las cuestiones 28 y 29 hacen referencia al siguiente enunciado:

Sean X e Y dos v.a. independientes con distribuciones respectivas $X \in N(m_X, \sigma_X^2 = 25)$ e $Y \in N(m_Y, \sigma_Y^2 = 36)$. Para contrastar $H_0 : m_X = m_Y$ frente a $H_1 : m_X \neq m_Y$ se toman en ambos colectivos sendas m.a.s. de tamaño 30 que proporcionan $\overline{x} = 82$ e $\overline{y} = 80$.

28. Un intervalo de confianza 90% para $(m_{X-}m_Y)$ es, aproximadamente:

(A) $(-0.34,\ 4.34)$ (B) $(-1.33,\ 5.33)$ (C) $(0.79,\ 4.79)$ (D) $(0.34,\ 4.34)$ (E) $(-0.79,\ 4.79)$

29. La decisión del contraste al nivel de significación del 10% será:

(A) No rechazar H_0 (B) - (C) Rechazar H_0 (D) - (E) -

30. Se sabe que el gasto en chucherías en una cierta población sigue una distribución normal y se sospecha que la varianza del gasto en tales productos efectuado por individuos de menos de 20 años, σ_1^2, es igual o mayor que la varianza del gasto efectuado por individuos de edad superior, σ_2^2. Para contrastar $H_0 : \sigma_1^2 \geq \sigma_2^2$ frente a $H_1 : \sigma_1^2 < \sigma_2^2$ se seleccionan muestras de tamaño igual a 31 en cada colectivo. Las m.a.s. proporcionan $s_1^2 = 62$ y $s_2^2 = 65$. La decisión del contraste, al nivel de significación del 10%, será:

(A) Rechazar H_0 (B) - (C) - (D) - (E) No rechazar H_0

PROBLEMAS (Duración: 75 minutos)

A. (10 puntos, 25 minutos)

Sea $X_1, \ldots, X_n$ una m.a.s. de tamaño n tomada de una v.a. X cuya función de densidad de probabilidad viene dada por:

$$f(x,\theta) = \frac{1}{24\theta^5}\, x^4\, e^{-\frac{x}{\theta}} \qquad x > 0, \qquad \theta > 0$$

Se sabe que $\mathrm{E}(X) = 5\theta$ y que $\mathrm{Var}(X) = 5\theta^2$.

i) Obtén, **detalladamente**, el estimador por el método de momentos para el parámetro θ.

ii) Obtén, **detalladamente**, el estimador por máxima verosimilitud para el parámetro θ.

iii) ¿Es el estimador por el método de momentos de θ insesgado?, ¿consistente? y ¿eficiente?

B. (10 puntos, 25 minutos)

Sea $X_1, \ldots, X_4$ una m.a.s. tomada de una población con distribución de Poisson de parámetro λ. Queremos contrastar la hipótesis nula $H_0 : \lambda = 1$ frente a la hipótesis alternativa $H_1 : \lambda = 2$.

i) Obtén la forma de la región crítica de mayor potencia para dicho contraste.

ii) Al nivel de significación del 10%, calcula la región crítica del contraste.

iii) ¿Cuál es la potencia del contraste?

C. (10 puntos, 25 minutos) En una determinada Universidad se desea contrastar si las notas de un curso de Estadística dependen de la carrera que se esté realizando. Para ello, se toma una muestra entre los alumnos que asisten a clase, clasificándolos por su nota final y su carrera, obteniendo los siguientes resultados:

	Psicología	Medicina	Farmacia	Totales
Sobresaliente	11	28	22	61
Notable	20	34	30	84
Aprobado	22	8	13	43
Suspenso	6	4	9	19
Totales	59	74	74	207

Con estos datos, realiza el contraste solicitado al nivel de significación del 5%.

SOLUCIONES DEL CUESTIONARIO (tipo 0)

1: C	11: D	21: C
2: C	12: C	22: C
3: D	13: A	23: B
4: E	14: A	24: A
5: E	15: B	25: C
6: E	16: B	26: E
7: B	17: A	27: A
8: A	18: C	28: A
9: C	19: B	29: A
10: D	20: A	30: E

SOLUCIONES DE LOS PROBLEMAS

Problema A

La función de densidad de la v.a. X es:

$$f(x,\theta) = \frac{1}{24\theta^5} x^4 e^{-\frac{1}{\theta}x} \qquad x > 0, \qquad \theta > 0$$

Sabemos que $\mathrm{E}(X) = 5\theta$ y $\mathrm{Var}(X) = 5\theta^2$.

i) Estimador por el método de momentos

Igualamos el primer momento poblacional al primer momento muestral. Es decir,

$$\alpha_1 = \mathrm{E}(X) = a_1 = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$$

De la v.a. X sabemos que $\mathrm{E}(X) = 5\theta$, por lo que tendremos que:

$$5\theta = \frac{\sum_{i=1}^{n} X_i}{n} = \overline{X}$$

y, por tanto,

$$\hat{\theta}_{MM} = \frac{\overline{X}}{5}$$

ii) Estimador máximo verosímil

La función de verosimilitud de la muestra es:

$$\begin{aligned} L(\theta) &= f(x_1;\theta) \ \dots \ f(x_n;\theta) = \\ &= \left(\frac{1}{24\theta^5} x_1^4 e^{-\frac{x_1}{\theta}} \right) \ \dots \ \left(\frac{1}{24\theta^5} x_n^4 e^{-\frac{x_n}{\theta}} \right) = \\ &= \frac{1}{24^n \theta^{5n}} \left(\prod_{i=1}^{n} x_i \right)^4 e^{-\frac{\sum_{i=1}^{n} x_i}{\theta}} \end{aligned}$$

Calculamos su logaritmo neperiano:

$$\ln L(\theta) = -n\ln(24) - 5n\ln\theta + \ln\left[\prod_{i=1}^{n} x_i\right]^4 - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

Si derivamos respecto de θ e igualamos a cero se tiene:

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \frac{-5n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = 0$$

de donde,

$$-5n\theta + \sum_{i=1}^{n} x_i = 0$$

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^{n} X_i}{5n} = \frac{\overline{X}}{5}$$

iii) Insesgadez:

Para ver si el estimador es insesgado hay que comprobar si $\mathrm{E}(\hat{\theta}_{MM}) = \theta$.

En este caso,

$$\mathrm{E}(\hat{\theta}_{MM}) = \mathrm{E}\left(\frac{1}{5}\ \overline{X}\right) = \frac{1}{5}\ \mathrm{E}(X) = \frac{1}{5}\ (5\theta) = \theta$$

Por lo tanto, sí es insesgado.

Consistencia

Para ver si el estimador es consistente calculamos la varianza del estimador.

$$\mathrm{Var}(\hat{\theta}_{MM}) = \mathrm{Var}\left(\frac{1}{5}\ \overline{X}\right) = \frac{1}{25}\ \frac{\mathrm{Var}(X)}{n} = \frac{1}{25}\ \frac{5\theta^2}{n} = \frac{\theta^2}{5n}$$

Dado que se trata de un estimador insesgado y cuya varianza tiende a 0 cuando n tiende a infinito, se cumplen las condiciones suficientes para la consistencia. Podemos afirmar que se trata de un estimador consistente para θ.

Eficiencia

Para ver si el estimador es eficiente hay que comprobar si la varianza del estimador coincide con la cota de Cramer-Rao para un estimador regular e insesgado de θ.

La cota es $L_c = \dfrac{1}{n\mathrm{E}\left[\dfrac{\partial \ln(f(x,\theta))}{\partial\theta}\right]^2}$

Para calcularla haremos:

$$\ln f(x,\theta) = -\ln(24) - 5\ln(\theta) + 4\ln x - \frac{x}{\theta}$$

$$\frac{\partial \ln f(x,\theta)}{\partial\theta} = -\frac{5}{\theta} + \frac{x}{\theta^2} = \frac{1}{\theta^2}(x - 5\theta)$$

$$\left[\frac{\partial \ln(f(x,\theta))}{\partial\theta}\right]^2 = \frac{1}{\theta^4}(x-5\theta)^2$$

$$\mathrm{E}\left[\frac{\partial \ln(f(x,\theta))}{\partial\theta}\right]^2 = \frac{1}{\theta^4}\mathrm{E}(X-5\theta)^2 = \frac{1}{\theta^4}\mathrm{Var}(X) = \frac{5\theta^2}{\theta^4} = \frac{5}{\theta^2}$$

$$n\mathrm{E}\left[\frac{\partial \ln(f(x,\theta))}{\partial\theta}\right]^2 = \frac{5n}{\theta^2}$$

y

$$L_c = \frac{\theta^2}{5n} = \mathrm{Var}(\hat{\theta}_{MV})$$

En consecuencia, $\hat{\theta}_{MM}$ es un estimador eficiente de θ.

Problema B

Sea $X_1, \ldots, X_n$ una m.a.s. de tamaño $n = 4$ tomada de una población con distribución de Poisson de parámetro λ. Queremos contrastar la hipótesis nula $H_0 : \lambda = 1$ frente a la hipótesis alternativa $H_1 : \lambda = 2$.

i) Para deducir la forma de región crítica de mayor potencia hacemos uso del Teorema de Neyman-Pearson. Así, las funciones de verosimilitud bajo las hipótesis nula y alternativa, estarán dadas por:

$$L(\vec{x}; \lambda_0) = L(\vec{x}; \lambda = 1) = \left(e^{-n}(1)^{\sum_{i=1}^n x_i}\right) / \prod_{i=1}^{n} x_i!,$$

y

$$L(\vec{x}; \lambda_1) = L(\vec{x}; \lambda = 2) = \left(e^{-2n}(2)^{\sum_{i=1}^n x_i}\right) / \prod_{i=1}^{n} x_i!,$$

respectivamente. Por tanto, si aplicamos el Teorema de Neyman-Pearson, tendremos que:

$$\frac{L(\vec{x}; \lambda_0)}{L(\vec{x}; \lambda_1)} = \frac{\left(e^{-n}(1)^{\sum_{i=1}^n x_i}\right) / \prod_{i=1}^n x_i!}{\left(e^{-2n}(2)^{\sum_{i=1}^n x_i}\right) / \prod_{i=1}^n x_i!} \leq K, \; K > 0$$

$$\Longrightarrow e^n \left(\frac{1}{2}\right)^{\sum_{i=1}^n x_i} \leq K \Longrightarrow \left(\frac{1}{2}\right)^{\sum_{i=1}^n x_i} \leq K_1$$

Tomando logaritmos neperianos, tenemos que:

$$\left(\sum_{i=1}^{n} x_i\right) \ln\left(\frac{1}{2}\right) \leq K_2, \; K_2 > 0$$

Ahora, como $0.50 < 1$, el logaritmo es negativo, por lo que la desigualdad cambia de sentido y tendremos que la regla de decisión será rechazar la hipótesis nula si $Z = \sum_{i=1}^n X_i \geq C$.

ii) Al nivel de significación $\alpha = 0.10$, teniendo en cuenta que $Z = \sum_{i=1}^n X_i \in \mathcal{P}(n\lambda)$ y que $n = 4$, tendremos que:

$$\alpha = 0.10 \geq P[Z \geq C | H_0] = P[Z \geq C | Z \in \mathcal{P}(4)] = 1 - F_Z(C - 1)$$

$$\Longrightarrow F_Z(C - 1) \geq 0.90 \Longrightarrow (C - 1) = 7 \Longrightarrow C = 8.$$

Es decir, rechazamos la hipótesis nula si $Z = \sum_{i=1}^n X_i \geq 8$.

iii) Para calcular la potencia tendremos que:

$$\text{Pot} = P[Z \geq 8 | H_1] = P[Z \geq 8 | Z \in \mathcal{P}(8)] = 1 - F_Z(7) = 1 - 0.452961 = 0.547039.$$

Problema C

Se trata de un **contraste de independencia.** Los datos son los que aparecen a continuación:

	Psicología (PSI)	Medicina (MED)	Farmacia (FAR)	Totales
Sobresaliente (SO)	11	28	22	61
Notable (N)	20	34	30	84
Aprobado (A)	22	8	13	43
Suspenso (S)	6	4	9	19
Totales	59	74	74	207

Las probabilidades $\hat{p}_{i.}$ y $\hat{p}_{.j}$ se estiman a partir de los datos. Así,

$$\hat{p}_{\mathrm{SO},\bullet} = \frac{61}{207} \qquad \hat{p}_{\mathrm{N},\bullet} = \frac{84}{207} \qquad \hat{p}_{\mathrm{A},\bullet} = \frac{43}{207} \qquad \hat{p}_{\mathrm{S},\bullet} = \frac{19}{207}$$

$$\hat{p}_{\bullet,\mathrm{PSI}} = \frac{59}{207} \qquad \hat{p}_{\bullet,\mathrm{MED}} = \frac{74}{207} \qquad \hat{p}_{\bullet,\mathrm{FAR}} = \frac{74}{207}$$

Construimos la tabla:

	n_{ij}	$\hat{p}_{i,j} = \hat{p}_{i,\bullet} \cdot \hat{p}_{\bullet,j}$	$n\hat{p}_{i,j}$	$\frac{(n_{ij}-n\hat{p}_{i,j})^2}{n\hat{p}_{i,j}}$
SO, PSI	11	$61 \cdot 59/207^2$	17.386	2.346
SO, MED	28	$61 \cdot 74/207^2$	21.807	1.759
SO, FAR	22	$61 \cdot 74/207^2$	21.807	0.002
N, PSI	20	$84 \cdot 59/207^2$	23.942	0.649
N, MED	34	$84 \cdot 74/207^2$	30.029	0.525
N, FAR	30	$84 \cdot 74/207^2$	30.029	0.000
A, PSI	22	$43 \cdot 59/207^2$	12.256	7.747
A, MED	8	$43 \cdot 74/207^2$	15.372	3.535
A, FAR	13	$43 \cdot 74/207^2$	15.372	0.366
S, PSI	6	$19 \cdot 59/207^2$	5.415	0.063
S, MED	4	$19 \cdot 74/207^2$	6.792	1.148
S, FAR	9	$19 \cdot 74/207^2$	6.792	0.718
	207	1	207	$z = 18.858$

Bajo la hipótesis nula de independencia, el estadístico $\sum_{i,j} \frac{(n_{ij}-n\hat{p}_{i,j})^2}{n\hat{p}_{i,j}} \sim \chi^2_{(k'-1)(k''-1)}$, donde k' es el número de categorías en que ha sido dividida la variable notas en el curso de Estadística ($k' = 4$) y k'' es el número de categorías en que ha sido dividida la variable carrera ($k'' = 3$).

La regla de decisión es rechazar la hipótesis nula al nivel de significación aproximado del 5% si:

$$z > \chi^2_{\overline{(4-1),(3-1)}|0.05} = \chi^2_{6|0.05}.$$

En este caso:

$$z = 18.858 > 12.60 = \chi^2_{6|0.05},$$

por lo que se rechaza la hipótesis nula de independencia a un nivel de significación del 5%. Es decir, se puede afirmar que las notas obtenidas en el curso de Estadística no son estadísticamente independientes de la carrera que se esté realizando.