

GUÍA DE EJERCICIOS PARA EL EXAMEN

1. Una autopista posee 4 pistas, y se desea investigar si los conductores tienen preferencia por alguna de ellas. Se observó la pista por la que transitaban 1000 automóviles, cuyos resultados se resumen en la siguiente tabla. ¿Hay suficiente evidencia para decir que algunas pistas son preferidas sobre otras? Use $\alpha = 0,05$.

Pista	Cantidad observada
1	294
2	276
3	238
4	192

2. La cantidad de accidentes sufridos por maquinistas de una cierta industria se observó durante un periodo de tiempo, cuyos resultados se muestran en la siguiente tabla. Realice un test a nivel 5% sobre la hipótesis de que los datos provienen de una distribución de Poisson.

Accidentes por maquinista	Cantidad de maquinistas
0	296
1	74
2	26
3	8
4	4
5	4
6	1
7	0
8	1

3. Se desea investigar la efectividad de una nueva vacuna. Se toma una muestra de 1000 personas, a algunas de las cuales se les da 1 dosis de la vacuna, a otras se les da 2 dosis, y al resto no se les vacuna. Pasadas 2 semanas se observa si las personas enfermaron. Se resume la información en la siguiente tabla:

	Sin vacuna	1 dosis	2 dosis
Enfermo	24	9	13
Sano	289	100	565

¿Hay suficiente evidencia para afirmar que la vacuna tiene un efecto observable?

4. En un estudio se clasificó a 81 televidentes en “audiencia de alta violencia” y “audiencia de baja violencia”, de acuerdo a los programas de televisión que ellos habitaban ver. Los resultados, segmentados por grupos de edad, se muestran en la siguiente tabla. ¿Hay suficiente evidencia para afirmar, al nivel 5%, que el tipo de audiencia

es independiente de la edad del televidente? Entregue el p -valor del test (o al menos una cota).

Violencia \ Edad	16-34	35-54	55 ó más
Baja	8	12	21
Alta	18	15	7

5. Consideramos dos indicadores demográficos para 20 países de América Latina: la esperanza de vida, denotada y , y la tasa de natalidad, denotada x . Se estudia la relación que existe entre estas variables, a partir del modelo

$$y \approx \beta_0 + \beta_1 x.$$

Se cuenta con una muestra de tamaño $n = 70$, con $\frac{1}{n} \sum x_i y_i = 24,8$, $\frac{1}{n} \sum x_i^2 = 0,123$, $\bar{x} = 0,305$, $\bar{y} = 78,3$. Estime los parámetros β_0 y β_1 en el modelo, de manera que se minimicen los errores al cuadrado. ¿Cuál es la predicción del modelo para $x = 0,7$?

6. Se dispone de una variable aleatoria Y , la cual se pretende explicar mediante una combinación lineal de potencias de una variable X , de la forma

$$Y \approx \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_k X^k.$$

Dada una muestra $(x_1, y_1), \dots, (x_n, y_n)$, explícite el vector $\mathbb{Y}$ de variables dependientes, el vector β de coeficientes de la regresión, y la matriz $\mathbb{X}$ con los datos de la variable independiente.

7. Se dispone de una muestra aleatoria simple $(X_1, Y_1), \dots, (X_n, Y_n)$ de datos provenientes del par de variables (X, Y) , donde $\mathbb{E}(X)$, $\mathbb{E}(Y)$ y $\text{var}(X) > 0$ son finitas.

- a) Argumente que $\frac{1}{n} \sum_{i=1}^n X_i Y_i$ y $\frac{1}{n} \sum_{i=1}^n X_i^2$ convergen casi seguramente a $\mathbb{E}(XY)$ y $\mathbb{E}(X^2)$, respectivamente, cuando $n \rightarrow \infty$.
- b) Se ajusta un modelo lineal de la forma $Y \approx a + bX$, el cual se resuelve mediante el criterio de minimizar la suma de residuos al cuadrado, con lo cual se obtienen las cantidades $\hat{a}_n$ y $\hat{b}_n$. Muestre que $\hat{b}_n \rightarrow \text{cov}(X, Y)/\text{var}(X)$ y $\hat{a}_n \rightarrow \mathbb{E}(Y) - \mathbb{E}(X)\text{cov}(X, Y)/\text{var}(X)$ casi seguramente cuando $n \rightarrow \infty$.