Laboratorio 12 - Inferencia Estadística

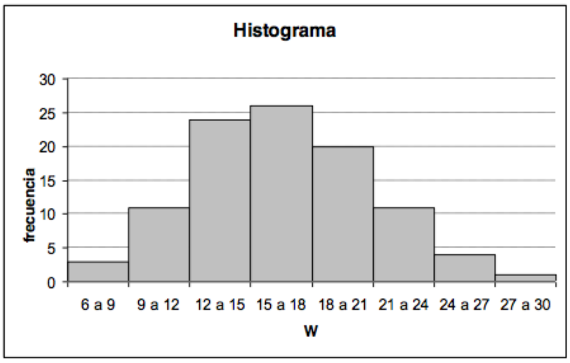
Pruebas de normalidad y el problema de predicción

Laboratorista: Héctor Lira Talancón

Ago-Dic 2017

1. Algunos analistas de mercado sostienen que últimamente el riesgo al invertir en el ramo de la construcción ha disminuido. Los resultados que se presentan a continuación corresponden a la variable W: rendimiento mensual (%), después de obtener una muestra aleatoria de este tipo de inversiones.

   Datos: $n = 100, X_{(1)} = 7.6125, X_{(n)} = 28.3334, Q_1 = 13.8849, Q_2 = 16.2717, Q_3 = 19.5843,$
   $s = 4.3248, C_A = 0.8178, C_K = 0.2719$

   Antes de realizar una prueba sobre el riesgo, es conveniente saber si la muestra que se tiene se puede considerar proveniente de una población normal. ¿Es razonable pensar que efectivamente el rendimiento sigue aproximadamente una distribución normal? Justificar la respuesta.

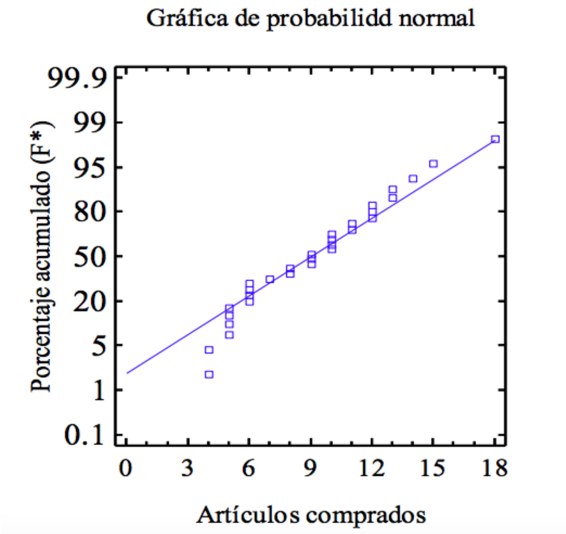
2. El gerente de una tienda de Wal-Mart estudia la cantidad de artículos que compran los consumidores en el horario de la tarde. A continuación se presenta la cantidad de artículos de una muestra aleatoria de 30 consumidores, organizados en una tabla de frecuencias.

Clase	Frecuencia
$[0, 3)$	0
$[3, 6)$	6
$[6, 9)$	7
$[9, 12)$	9
$[12, 15)$	6
$[15, 18)$	1
$[18, 21)$	1

   a) Calcula la media, mediana y desviación estándar con los datos agrupados.

b) Construye la distribución de frecuencia relativa acumulada de los datos.

c) A partir de los resultados del inciso b) y la gráfica de probabilidad Normal que se muestra a continuación, ¿se podría concluir que los datos provienen de una distribución Normal?



3. De acuerdo a lo visto en clase, el criterio para obtener la predicción $\hat{Y}$ del valor de una variable aleatoria Y es,

   a) Minimizar el valor esperado del cuadrado del error de predicción.

   b) Maximizar la probabilidad de que Y sea igual a $\hat{Y}$.

   c) Que la varianza de $\hat{Y}$ sea cero.

   d) Ninguna de las anteriores.

4. Se tienen X e Y dos variables aleatorias con varianza finita. Supongamos que $(\hat{\alpha}, \hat{\beta})$ minimiza a $G(a,b) = E[(Y - a - bX)^2]$. Considera las siguientes afirmaciones:

   A: $G(\hat{\alpha}, \hat{\beta}) \leq G(E(Y), 0)$

   B : $G(\hat{\alpha}, \hat{\beta}) < E[\{Y - E(Y|X)\}^2]$

   Entonces:

   a) A es falsa y B es verdadera.

   b) A y B son verdaderas.

   c) A es verdadera y B es falsa.

   d) Ninguna de las anteriores.

5. Sean X e Y variables aleatorias con varianza finita. Definimos $EC(a,b) = E[(Y - a - bX)^2]$. Considera las siguientes afirmaciones:

   A: $EC(E[Y], 0) > EC(x, 0)$ si $x \neq E(Y)$

   B: Supongamos que (β_0, β_1) minimizan a $EC(a,b)$, entonces $EC(E[Y], 0) \geq EC(\beta_0, \beta_1)$.

Entonces,

a) A y B son falsas.

b) A es falsa y B verdadera.

c) A y B son verdaderas.

d) Ninguna de las anteriores.

6. Si se desea predecir el valor de una variable aleatoria Y con base en lo observado para otra variable aleatoria X, entonces es cierto que:

a) $E(X|Y)$ produce el mejor predictor de Y, para cualquier elección de X.

b) Mientras mayor sea la covarianza entre Y y X, menor será la varianza de Y, y por ello habrá menos incertidumbre.

c) El mejor predictor lineal de Y minimiza el error cuadrático medio entre Y y su esperanza condicional dado X.

d) Para que la esperanza condicional sea el mejor predictor de Y, se requiere que X se comporte en forma lineal.

7. Suponga que X e Y son variables aleatorias tale que $f_{X,Y}(x,y) = \frac{1}{x}\exp\{-\frac{x^2+y}{x}\}\mathbb{1}_{(0,\infty)}(x)\mathbb{1}_{(0,\infty)}(y)$. Para estas variables se puede demostrar que $f_{Y|X}(y|x) = \frac{1}{x}\exp\{-\frac{y}{x}\}\mathbb{1}_{(0,\infty)}(y)$. Analice las siguientes afirmaciones para elegir la opción correcta:

A: El predictor mínimo cuadrático de Y, denotado por $\hat{Y}$, se obtiene minimizando $E[(Y-\hat{Y})^2]$.

B: Si se quiere predecir Y a través de una constante, entonces su predictor mínimo cuadrático es $E[Y]$.

C: Si se quiere predecir a Y a través de X, y se sabe que $X = x_0$, entonces su predictor mínimo cuadrático es x_0.

Entonces,

a) Solo A y B son verdaderas.

b) Solo B y C son verdaderas.

c) Solo A y C son verdaderas.

d) A, B y C son verdaderas.

8. Suponga que Y es una variable aleatoria y que su predictor es de la forma $\hat{Y} = \alpha X + \beta$. Demuestre que $\hat{Y} = E[Y] + \frac{Cov[X,Y]}{Var[X]}(X - E[X])$ es su predictor mínimo cuadrático. ¿Es $\hat{Y}$ un predictor insesgado?

9. Suponga que Y es una variable aleatoria y que su predictor es $\hat{Y} = g(X)$, donde g es una función arbitraria. Demuestre que $E\Big[\Big(Y - g(X)\Big)^2\Big] \geq E\Big[\Big(Y - E[Y|X]\Big)^2\Big]$.

10. Si $Z = h(X,Y)$, demuestre que $E[Z] = E_X\big[E[Z|X]\big]$.

11. Demuestre que $Var[Y] = E\big[Var[Y|X]\big] + Var\big[E[Y|X]\big]$.

12. Si $f_{X,Y}(x,y) = 2x\mathbb{1}_{(x,1)}(y)\mathbb{1}_{(0,1)}(x)$. Encuentre el mejor predictor mínimo cuadrático de Y.

13. Si $f_{X,Y}(x,y) = e^{-y}\mathbb{1}_{(0,y)}(x)\mathbb{1}_{(0,\infty)}(y)$.
    a) Encuentre el mejor predictor mínimo cuadrático de Y.
    b) Calcule la varianza condicional.