Universidad De Atacama

Facultad De Ingeniería / Departamento De Matemática

ESTADÍSTICA Y PROBABILIDADES

PAUTA DE CORRECCIÓN PRUEBA PARCIAL Nº3

Profesor: Hugo S. Salinas. Primer Semestre 2012

1. El nivel de llenado de unas botellas de bebidas gaseosas tiene una distribución normal con media 2 litros y desviación estándar 0.06 litros. Las botellas que contienen menos de 95 % del contenido neto anunciado pueden causar una multa al fabricante por parte del SERNAC, mientras que las botellas que tienen un contenido neto mayor que 2.1 litros pueden provocar un derrame del exceso al abrirlas.

   a) ¿Cuál es la probabilidad de que le pongan una multa al fabricante, si se selecciona al azar una botella de la producción?
   Solución:
   Sea X : nivel del llenado de las botellas gaseosas en litros, donde $X \sim N(2, 0.06)$. De acuerdo al enunciado, si las botellas tienen menos del 95 % del contenido neto anunciado (es decir 1.9 litros) pueden causar una multa. Por lo tanto:

   $$P(X < 1.9) = P(Z < (1.9 - 2)/0.06)) = P(Z < -1.67) = 0.0475$$

   Es decir, hay un 4.75 % de posibilidad que le pongan una multa al fabricante, donde $Z \sim N(0, 1)$.

   (5 ptos.)

   b) ¿Qué proporción de las botellas pueden provocar un derrame al abrirlas?
   Solución:
   De acuerdo al enunciado, se pide:

   $$P(X > 2.1) = P(Z > (2.1-2)/0.06) = P(Z > 1.67) = 1-P(Z < 1.67) = 1-0.9525 = 0.0475$$

   Es decir, aproximadamente 5 de cada 100 botellas pueden provocar derrames al abrirlas.

   (5 ptos.)

   c) ¿Qué cantidad mínima de refresco se espera que contenga 99 % de las botellas?
   Solución:
   Sea M la cantidad mínima para que las botellas contengan el 99 % de refresco. De acuerdo a esto se necesita resolver:

   $$P(X > M) = 0.99 \iff P(Z > (M-2)/0.06) = 0.99 \iff P(Z < (M-2)/0.06) = 0.01$$

   Por lo tanto:

   $$\frac{M-2}{0.06} = -2.32 \iff M - 2 = -0.1392 \iff M = 1.8608$$

   Por lo tanto, se espera que tengan 1.8608 litros como cantidad mínima.

   (5 ptos.)

d) En un día se llenan 100 botellas ¿cuál es la probabilidad de que haya en un día más de 4 botellas que puedan provocar un derrame al abrirlas?
<u>**Solución**</u>:
Sea Y : número de botellas que, en un día, puedan provocar un derrame al abrirlas de un total de 100. De acuerdo a esto, se trata de una variable aleatoria binomial con $n = 100$ y $p = P(X > 2.1) = 0.0475$. Se pide calcular:

$$\begin{aligned} P(Y > 4) &= 1 - P(Y \leq 4) = 1 - \sum_{y=0}^{4} \binom{100}{y} (0.0475)^y (0.9525)^{100-y} \\ &= 1 - (0.00770 + 0.03840 + 0.09479 + 0.15442 + 0.18674) \\ &= 0.51795 \end{aligned}$$

(5 ptos.)

e) Utilizando el apartado anterior, ¿cuál es, en un mes de 30 días, el número medio de días en los que se producen más de 4 botellas que puedan provocar un derrame al abrirlas?
<u>**Solución**</u>:
De acuerdo a la variable aleatoria del item anterior, se trata de calcular el valor esperado. Como se sabe $E(Y) = np$, entonces lo que se pide es $E(Y) = 30 \times 0.51795 = 15.5385$. Por lo tanto, hay 16 días en promedio, en los que se producen más de 4 botellas que puedan provocar un derrame al abrirlas.

(5 ptos.)

2. El tiempo T en segundos que tarda un PC en conectarse a un servidor durante un día laboral sigue una distribución de Weibull con parámetros $\alpha = 1/4$ y $\beta = 0.6$, mientras que en un día de fin de semana el tiempo es una variable aleatoria Exponencial con $\beta = 0.24$.

a) Calcular el tiempo medio que tarda un PC en conectarse al servidor en ambos tipos de días.
<u>**Solución**</u>:
Primero llamemos T_L al tiempo, en segundos, que tarda el PC en conectarse a un servidor durante un día laboral y T_F al tiempo, en segundos, que tarda un PC en un día de fin de semana. Del enunciado se sabe que $T_L \sim W(\alpha = 1/4, \beta = 0.6)$ y $T_F \sim \mathcal{E}(\beta = 0.24)$. Por lo tanto, utilizando el formulario se tiene que:

$$E(T_L) = 0.6\Gamma\left(1 + \frac{1}{1/4}\right) = 0.6\Gamma(5) = 4! \times 0.6 = 14.4$$

$$E(T_F) = 0.24$$

(6 ptos.)

b) Calcular la probabilidad, para ambos tipos de días, de que un PC tarde menos de 10 segundos en realizar la conexión.
<u>**Solución**</u>:
Esto se consigue directamente del formulario, es decir se pide

$$P(T_L < 10) = F_{T_L}(10) = 1 - e^{-(\frac{10}{0.6})^{1/4}} = 0.87$$

$$P(T_F < 10) = F_{T_F}(10) = 1 - e^{-\frac{10}{0.24}} = 1$$

(6 ptos.)

c) Si un PC lleva 5 segundos esperando a que se realice la conexión con el servidor durante un día laboral. ¿Cuál es la probabilidad de que la conexión se demore aun 10 segundos más?
Solución:
En este caso se pide calcular $P(T_L > 15|T_L > 5)$. Utilizando la propiedad de la probabilidad condicional se tiene que

$$\begin{aligned} P(T_L > 15|T_L > 5) &= \frac{P(\{T_L > 15\} \cap \{T_L > 5\})}{P(T_L > 5)} = \frac{P(T_L > 15)}{P(T_L > 5)} \\ &= \frac{1 - F_{T_L}(15)}{1 - F_{T_L}(5)} = \frac{e^{-(\frac{15}{0.6})^{1/4}}}{e^{-(\frac{5}{0.6})^{1/4}}} = 0.58 \end{aligned}$$

(6 ptos.)

d) Probar que si X es una Weibull de parámetro $\alpha = 1$ se tiene que

$$P(X > t_0 + t|X > t_0) = P(X > t)$$

para todo $\beta > 0$.
Solución:
Se sabe que si $\alpha = 1$ entonces la variable aleatoria X es una exponencial de parámetro $\beta > 0$, por lo tanto:

$$\begin{aligned} P(X > t_0 + t|X > t_0) &= \frac{P(\{X > t_0 + t\} \cap \{X > t_0\})}{P(X > t_0)} = \frac{P(X > t_0 + t)}{P(X > t_0)} \\ &= \frac{1 - F_X(t_0 + t)}{1 - F_X(t_0)} = \frac{e^{-\frac{t_0+t}{\beta}}}{e^{-\frac{t_0}{\beta}}} = \frac{e^{-\frac{t_0}{\beta}} e^{-\frac{t}{\beta}}}{e^{-\frac{t_0}{\beta}}} \\ &= e^{-\frac{t}{\beta}} = 1 - F_X(t) = P(X > t). \end{aligned}$$

(7 ptos.)

3. En un cierto periódico, la longitud en pulgadas de las columnas de anuncios clasificados que aparecen los lunes tiene una distribución aproximadamente normal, con una media de 327 pulgadas y una desviación estándar de 34 pulgadas. Considerar las medidas de 10 lunes consecutivos como una muestra aleatoria.

a) Calcular el valor esperado y la desviación estándar de la longitud total (en pulgadas) de las columnas de anuncios clasificados para 10 lunes.
Solución:
Sea X: longitud de las columnas de los anuncios, donde $X \sim N(\mu_X = 327, \sigma_X^2 = 34^2)$. Por otro lado, sea $T = \sum_{i=1}^{10} x_i$ la longitud total de las columnas para 10 lunes. Entonces $T \sim N(\mu = 10 \times 327, \sigma^2 = 10 \times 34^2)$, de aquí es fácil ver que el valor esperado de T es 3270 y la desviación estándar es $\sqrt{10 \times 34^2} = 107.52$.

(9 ptos.)

b) Calcular la probabilidad de que el total se encuentre entre 3150 y 3390 pulgadas.
Solución:

$$\begin{aligned} P(3150 < T < 3390) &= P(T < 3390) - P(T < 3150) \\ &= P(Z < (3390 - 3270)/107.52) - P(Z < (3150 - 3270)/107.52) \\ &= P(Z < 1.12) - P(Z < -1.12) \\ &= 0.8686 - 0.1314 = 0.7372. \end{aligned}$$

(8 ptos.)

c) Calcular la probabilidad de que la longitud promedio de las columnas para cada lunes se encuentre entre 314 y 339.

Solución:

Sea $\overline{X}$ la longitud promedio para cada lunes, donde $\overline{X} \sim N(\mu_X, \sigma_X^2/n)$ cuando $n \to \infty$, es decir,

$$\overline{X} \sim N\left(327, \frac{34^2}{10}\right), \qquad n \to \infty$$

Por lo tanto,

$$\begin{aligned} P(314 < \overline{X} < 339) &= P(\overline{X} < 339) - P(\overline{X} < 314) \\ &= P(Z < (339 - 327)/10.75) - P(Z < (3150 - 3270)/10.75) \\ &= P(Z < 1.12) - P(Z < -1.21) \\ &= 0.8686 - 0.1131 = 0.7555. \end{aligned}$$

(8 ptos.)

4. RESOLVER:

a) EL supervisor del proceso de empacado de té en sobres, selecciona una muestra aleatoria de 12 sobres y mide el peso neto de cada uno, obteniendo la siguiente información: 15.7 - 15.8 - 15.8 - 15.9 - 15.9 - 16.0 - 16.0 - 16.0 - 16.1 - 16.1 - 16.1 - 16.2. Si el peso neto de los sobres sigue una distribución normal, determinar un intervalo al 90 % de confianza para la media poblacional de las bolas de té. **Nota**: $\overline{X^2} = 254.955$

Solución:

Buscamos un intervalo de confianza para la media cuando no se conoce la varianza (no se dice algo sobre la varianza). Luego

$$\mu \in \left(\overline{X} \pm t_{\frac{\alpha}{2};n-1} \frac{S}{\sqrt{n}}\right)$$

Como en el enunciado nos dicen que la confianza es del 90 %, entonces $1 - \alpha = 0.9 \Leftrightarrow \alpha/2 = 0.05$. De la tabla de la t de Student calculamos el percentil $t_{\alpha/2}$ con $n - 1 = 11$ grados de libertad. Luego $t_{\alpha/2,11} = 1.796$.

Ahora calculamos la media y la varianza muestral:

$$\begin{aligned} \overline{X} &= \frac{\sum_i x_i}{n} = \frac{15.7 + 15.8 + \ldots + 16.2}{12} = 15.96667 \\ S^2 &= \frac{n}{n-1}[\overline{X^2} - (\overline{X})^2] = \frac{12}{11}[254.955 - (15.96667)^2] = 0.02231 \end{aligned}$$

De aquí $S = \sqrt{0.02231} = 0.15$ y finalmente:

$$\begin{aligned} \mu &\in \left(15.97 \pm 1.796 \frac{0.15}{\sqrt{12}}\right) \\ \mu &\in (15.8922,\ 16.0478) \end{aligned}$$

(6 ptos.)

b) En un país oriental se desea estudiar la variable altura de los individuos. Para esto se realizó un estudio piloto con una muestra de personas la cual arrojó una media de 170 cm. Calcular el tamaño que debería tener la muestra para obtener un intervalo de confianza al 99 % con una precisión (error) de un centímetro. **Nota**: Usar $\sigma = 10$ y asumir que la altura de los individuos distribuye normal.

<u>**Solución**</u>:

Tenemos que $\mu = 170$. Además sabemos que $\sigma = 10$. Entonces estamos trabajando con el siguiente intervalo para μ. Como nos piden un 99 % de confianza, tenemos que $\alpha/2 = 0.005$.

$$\begin{aligned} \mu &\in \left(\overline{X} \pm z_{\alpha/2}\,\frac{\sigma}{\sqrt{n}}\right) \\ \mu &\in \left(170 \pm z_{0.005}\,\frac{10}{\sqrt{n}}\right) \\ \mu &\in \left(\overline{X} \pm \underbrace{2.58\,\frac{10}{\sqrt{n}}}_{\text{error}}\right) \end{aligned}$$

Queremos un error de un centímetro, es decir, $2.58\,\frac{10}{\sqrt{n}} = 1 \Leftrightarrow 25.8 = \sqrt{n}$. Luego $n = 665.64$, pero como el tamaño de muestra debe ser entero, debemos hacer una aproximación, por lo tanto $n = 666$.

(6 ptos.)

c) Un directivo de cierta empresa ha comprobado que los resultados obtenidos en los test de aptitud por los solicitantes de un determinado puesto de trabajo sigue una distribución normal con una desviación estándar de 1.42 puntos. Las calificaciones de nueve test están dadas por: 188 - 190 - 185 - 187 - 188 - 187 - 188 - 189 - 189. Calcular un intervalo de confianza del 90 % para la calificación media poblacional del grupo de solicitantes actual.

<u>**Solución**</u>:

Primero calculamos la media de los datos:

$$\overline{X} = \frac{\sum X_i}{n} = \frac{188 + 190 + \ldots + 189 + 189}{9} = 187.9$$

El intervalo de confianza para la media, con un 90 % de confianza está dado por:

$$\mu \in \left(187.9 \pm 1.645\,\frac{1.42}{\sqrt{9}}\right)$$

Puesto que al 90 % de confianza $\alpha = 0.1$, entonces $z_{\alpha/2} = z_{0.05} = 1.645$ según tabla, luego:

$$\mu \in (187.12;\ 188.68)$$

(6 ptos.)

d) A partir de los datos anteriores (item c)), un estadístico calcula para la media poblacional un intervalo de confianza que va desde 187.8 a 188.0 puntos. Calcular el nivel de confiabilidad ($(1-\alpha) \times 100\,\%$) tomado por el estadístico.

<u>**Solución**</u>:

Tenemos que el intervalo de confianza calculado por la persona es:

$$\mu \in (187.8;\ 188.00)$$

Luego debemos resolver:

$$187.9 + z_{\alpha/2}\frac{1.42}{\sqrt{9}} = 188 \Longleftrightarrow z_{\alpha/2}\,1.42 = 0.3 \Longleftrightarrow z_{\alpha/2} = 0.21$$

Por lo tanto $1 - \frac{\alpha}{2} = 0.5832 \Longleftrightarrow \alpha = 0.8336$. Luego el nivel de confianza tomado por la persona es de un $(1-\alpha)\times 100\,\% = 16.64\,\%$.

(7 ptos.)