LEC/LADE/LECD/LADED
CURSO 2007/08
SOLUCIONES HOJA DE PROBLEMAS 1
DISTRIBUCIONES EN EL MUESTREO Y ESTIMACIÓN PUNTUAL

1.

a) (1 punto)

$$E(\bar{X}) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu = 22$$

Por ser las muestras, independientes:

$$V(\bar{X}) = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V\left[X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n} = \frac{9}{10} = .9$$

b) (1 punto) Sabemos que la cuasivarianza es un estimador insesgado para la varianza, por tanto,

$$E(s^2) = \sigma^2 = 9$$

Además, $\hat{\sigma}^2 = \frac{n-1}{n}s^2$, por tanto,

$$E(\hat{\sigma}^2) = \frac{n-1}{n}E(s^2) = \frac{9}{10}9 = 8.1$$

No se puede obtener el mismo valor puesto que la cuasivarianza es un estimador insesgado y la varianza no.

c) (i) (0.5 puntos) Sabemos que $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}}) = N(22, \frac{3}{\sqrt{100}})$(distribución de probabilidad).

$$P(\bar{X} < 25) = P(\frac{\bar{X}-22}{0.3} < \frac{25-22}{0.3}) = P(Z < 10) \cong 1$$

donde $Z \sim N(0,1)$.

(ii) (0.25 puntos) Debemos calcular $P(X < 25)$, donde $X \sim N(22, 3)$.

$$P(X < 25) = P(\frac{X-22}{3} < \frac{25-22}{3}) = P(Z < 1) = 0.8413$$

(iii) (0.25 puntos) En cuanto a la cuasidesviación típica sabemos que $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$. Entonces,

$$P(s > 4) = P(s^2 > 16) = P(\frac{(n-1)s^2}{\sigma^2} > \frac{(n-1)16}{\sigma^2}) = P(\chi^2_{99} > \frac{9916}{9}) = P(\chi^2_{99} > 176) \approx 0$$

(Aproximamos por una χ^2_{100} y además $P(\chi^2_{100} > 140.2) = .005$).

2.

a) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, donde S es la cuasi-desviación típica muestral.

$P(S < 1.99) = P\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(n-1)\times 1.99^2}{\sigma^2}\right) = P\left(\chi^2_{15} < \frac{15\times 1.99^2}{3.37^2}\right) = P\left(\chi^2_{15} < 5.23\right) = 0.01.$

b) $P(S > 2.89) = 1 - P(S \leq 2.89) = 1 - P\left(\frac{(n-1)S^2}{\sigma^2} \leq \frac{(n-1)\times 2.89^2}{\sigma^2}\right) =$

$$= 1 - P\left(\chi^2_{15} \leq \frac{15\times 2.89^2}{3.37^2}\right) = 1 - P\left(\chi^2_{15} \leq 11.03\right) = 1 - 0.25 = 0.75.$$

3. **a)** $E\left[\overline{X}\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{16}\sum_{i=1}^{16}(3.5) = 3.5.$

$V\left[\overline{X}\right] = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V\left[X_i\right] = \frac{1}{16^2}\sum_{i=1}^{16}(0.25) = 0.0156.$

$\overline{X} \sim N(3.5, 0.0156)$

En este caso se trata de una distribución exacta pues la suma de variables aleatorias normales es una variable aleatoria normal.

b) $P\left(\overline{X} > 3.7\right) = 1 - P\left(\overline{X} \leq 3.7\right) = 1 - P\left(\frac{\overline{X}-3.5}{0.1248} \leq \frac{3.7-3.5}{0.1248}\right) = 1 - P(N(0,1) \leq 1.6) =$

$$= 1 - 0.9452 = 0.0548.$$

c) $P\left(3.34 < \overline{X} < 3.36\right) = P\left(\overline{X} < 3.36\right) - P\left(\overline{X} < 3.34\right) =$

$$= P\left(N(0,1) < \frac{3.36-3.5}{0.1248}\right) - P\left(N(0,1) < \frac{3.34-3.5}{0.1248}\right) =$$

$$= P(N(0,1) < -1.121) - P(N(0,1) < -1.28) =$$

$$= 1 - P(N(0,1) < 1.121) - 1 + P(N(0,1) < 1.28) =$$

$$= -0.8686 + 0.8957 = 0.0311.$$

4.

(a) $E(\overline{X} - \overline{Y}) = E(\overline{X}) - E(\overline{Y}) = \mu_x - \mu_y.$

1. V$(\overline{X}-\overline{Y}) = V(\overline{X})+V(\overline{Y}) = \frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}$.Por ser las muestras independientes y $V(\overline{X}) = \frac{\sigma_x^2}{n_x}$ y $V(\overline{Y}) = \frac{\sigma_y^2}{n_y}$.

   $\overline{X}-\overline{Y} \sim N(\mu_x-\mu_y; \frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y})$ dado que $X \sim N(\mu_x;\sigma_x^2) \Rightarrow \overline{X} \sim N(\mu_x; \frac{\sigma_x^2}{n_x})$ y que $Y \sim N(\mu_y;\sigma_y^2) \Rightarrow \overline{Y} \sim N(\mu_y; \frac{\sigma_y^2}{n_y})$.

(b) Empleando la distribución en el muestreo de $\overline{X}-\overline{Y}$ tenemos que:

P$(\overline{X}-\overline{Y} > \mu_x-\mu_y) = P(\frac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}} > \frac{\mu_x-\mu_y-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x}+\frac{\sigma_y^2}{n_y}}}) =$

$P(z > 0) = 0,5$

5. Sea X el tiempo en acabar la tarea.

$$\mu = 50, \qquad \sigma = 8, \qquad n = 60$$

a) Queremos calcular $\mathbf{P(}\overline{X} \geq 52)$.

Como el tamaño muestral es suficientemente grande,

P$(\frac{(\overline{X}-50)\times\sqrt{60}}{8} \geq \frac{(52-50)\times\sqrt{60}}{8}) = P(Z \geq 1.9364) \simeq 0.0266$

b) La probabilidad es muy pequeña, luego sí tiene razones para pensar que ha aumentado el tiempo.

6.

a) $\sigma\left(\overline{X}\right) = \frac{3000}{\sqrt{100}} = 300.$

b) $\overline{X} \sim N\left(\mu, \sigma^2 = 300^2\right).$

$$P\left(\overline{X}-\mu > 200\right) = P\left(\frac{\overline{X}-\mu}{300} > \frac{200}{300}\right) = P\left(N\left(0,1\right) > \frac{2}{3}\right) = 0.2514.$$

c) $P\left(\overline{X} \leq \mu-300\right) = P\left(\frac{\overline{X}-\mu}{300} \leq -1\right) = P\left(N\left(0,1\right) \leq -1\right) = 0.1587.$

d) $P\left(\left|\overline{X}-\mu\right| > 400\right) = P\left(\frac{\left|\overline{X}-\mu\right|}{300} > \frac{4}{3}\right) = P\left(\left|N\left(0,1\right)\right| > \frac{4}{3}\right) =$

$$= 1-P\left(\left|N\left(0,1\right)\right| \leq \frac{4}{3}\right) = 1-P\left(-\frac{4}{3} \leq N\left(0,1\right) \leq \frac{4}{3}\right) =$$

$$= 1 - \left[P\left(N\left(0,1\right) \le \tfrac{4}{3}\right) - P\left(N\left(0,1\right) \le -\tfrac{4}{3}\right)\right] =$$

$$= 1 - (0.9082 - 0.0918) = 1 - 0.8164 = 0.1836.$$

7.

a) $E\left[\bar{X}\right] = E\left[\frac{{}^{n}_{i=1}X_i}{n}\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{n}n\mu = \mu$

$\sigma^2_{\bar{X}} = V\left[\bar{X}\right] = V\left[\frac{{}^{n}_{i=1}X_i}{n}\right] = \frac{1}{n^2}V\left[\sum_{i=1}^{n} X_i\right] = \{indep\} = \frac{1}{n^2}{}^{n}_{i=1}V\left[X_i\right] = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$

$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$

b) b.1) $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} = \frac{3{,}5}{\sqrt{16}} = 0{,}875$

b.2) $P(\bar{x} \le 10) = P(z \le \frac{10-12{,}1}{0{,}875}) = P(z \le -2{,}4) = 0.0082$

b.3) $P(s > 4{,}52) = P(\chi^2_{16-1} > \frac{(16-1)(4{,}52)^2}{(3{,}5)^2}) =$
$= P(\chi^2_{15} > 25) = 0{,}05$ (en este caso se ha tomado la desviación típica corregida). Como el ejercicio no especificaba qué desviación debía usarse, ambas contestaciones serán dadas por válidas (si no se supone desviación corregida en lugar de $\frac{(16-1)(4{,}52)^2}{(3{,}5)^2}$ habrá de usarse $\frac{(16)(4{,}52)^2}{(3{,}5)^2}$)

8.

a) (1 punto) $P(\overline{X} \ge 25) = P(\frac{\overline{X}-20}{3/\sqrt{100}} \ge \frac{25-20}{3/\sqrt{100}}) = P(N(0,1) \ge 16{,}66) \simeq 0.$
$P(\overline{X} = 21) = 0$ por ser variable aleatoria continua

b) (0.5 puntos) $p = 0{,}10 \Longrightarrow \frac{\widehat{p}_X - p}{\sqrt{\frac{p_X q_X}{n}}} N(0,1)$ si n suficientemente grande $\Longrightarrow$
$P(\widehat{p}_X < 0{,}08) = P(\frac{\widehat{p}_X - p}{\sqrt{\frac{p_X q_X}{n}}} < \frac{0{,}08-0{,}10}{\sqrt{\frac{0{,}10*0{,}90}{100}}})P(N(0,1) < -0{,}66) = 0{,}2546.$

c) (1 punto) $EFR(\overline{X}/\widehat{\mu}) = \frac{VAR(\widehat{\mu})}{VAR(\overline{X})} = \frac{\sigma^2}{2} / \frac{\sigma^2}{100} = 50 \Longrightarrow \overline{X}$ es 50 veces más eficiente que $\widehat{\mu}$