Tema 1. Distribucions en el mostreig

1. En un hospital pediàtric es considera que l'alçada, X, de la població de nens de 10 anys segueix una distribució normal de mitjana 135 cm i amb desviació estàndar 8cm.

   a) Calculeu $\mathrm{P}(X \geq 145)$.

   b) Determineu c tal que $\mathrm{P}(X \leq c) = 0{,}975$.

   c) Més que en la v.a. X, estem interessats en la mitjana mostral $\bar{X}$. Si considerem mostres de 16 nens, calculeu $\mathrm{P}(\bar{X} \geq 145)$.

   d) Determineu c tal que $\mathrm{P}(\bar{X} \leq c) = 0{,}975$.

2. La companyia de paqueteria PETEX està especialitzada en trameses internacionals per vaixell i té estimat que la mitjana del pes dels paquets que envia és de 8 Kg amb desviació estàndar de 4 Kg. La capacitat dels containers que habitualment usa és de una tona (1000 Kg). La companyia ha rebut 100 paquets de 100 clients diferents, quina és la probabilitat que els pugui empaquetar en una mateix container? Quina relació heu de suposar entre els clients per poder resoldre el problema?

3. Si tirem 1000 vegades una moneda perfecta, és a dir, $P(\text{cara}) = P(\text{creu})$:

   a) Quina és la probabilitat que el nombre de cares estigui comprés entre 490 i 510.

   b) Quin és l'interval centrat en 500 tal que el nombre de cares observat en els llançaments pertany a l'interval amb probabilitat 0,95?

4. El Servei de Parcs i Jardins d'un Ajuntament vol comprar a un centre de jardineria lots de 20 plantes però està interessat amb que l'alçada de les plantes, una vegada crescudes, sigui molt semblant. El centre de jardineria garanteix que l'alçada de les plantes segueix una distribució normal de mitjana 70 cm i desviació estandar 5 cm. Quina és la probabilitat que la variància mostral d'un lot sigui inferior a 10 cm^2?

5. Un de cada 400 nens que neixen en un hospital esta afectat per una malaltia congènita. Es demana:

   a) Quina és la probabilitat de trobar entre 3 i 5 nens afectats d'entre un total de 800?

   b) Sabent que d'un total de 600 nens com a mínim hi ha dos afectats quina és la probabilitat de trobar-ne tres?

c) Quina és la probabilitat de tenir més de 12 nens afectats d'un total de 4000?

6. Una agència de publicitat ha posat en marxa una campanya de promoció d'un nou producte. Al final de la campanya, l'agència sosté que un 25 % dels consumidors coneixen el producte. Per verificar aquesta suposició, el productor selecciona aleatòriament 1000 consumidors i observa que 232 si que coneixen el producte. Si certament el 25 % dels consumidors coneixen el producte, quina és la probabilitat de que, en una mostra de 1000 consumidors, hi hagi almenys 232 consumidors que coneguin el producte?

7. La durada d'una bombeta és una variable aleatòria amb distribució normal de mitjana 100 hores i de desviació estàndar 10 hores.

   a) Calcula la probabilitat que la mitjana d'un lot de 25 bombetes sigui superior a 95 hores.

   b) De quina mida hauria de ser un lot per tal que la diferencia entre la mitjana mostral i la mitjana sigui inferior a 5 hores amb probabilitat del 95 %?

8. La companyia area FLYFLY afirma que el nombre de passatgers que tenen reserva i no es presenten és del 6 %. Si en un vol de Barcelona a Estocolm que pot admetre 250 passatgers hi ha 260 reserves, calcula la probabilitat que la companyia pugui acomodar a tots els passatgers amb reserva que apareguin per embarcar.

9. Suposeu que hem de prendre una mostra d'una distribució normal amb esperança μ desconeguda i desviació estàndar 2. Calculeu la grandària de la mostra necessària, en cadascuna de les següents situacions, per tal que per a cada possible valor de μ:

   a) Tinguem una probabilitat superior o igual a 0.95 que la "desviació" entre la mitjana mostral i la mitjana poblacional sigui a tot estirar una dècima, és a dir:

   $$\mathrm{P}(\mid \bar{X}_n - \mu \mid \leq 0{,}1) \geq 0{,}95.$$

   b) En terme mig (esperança) la "desviació" quadràtica entre la mitjana mostral i la mitjana poblacional sigui a tot estirar una dècima, és a dir:

   $$E(\mid \bar{X}_n - \mu \mid^2) \leq 0{,}1.$$

Exercicis resolts

Exercici 1

Tenim la v.a. X que descriu l'alçada en cm i que segueix la distribució $N(\mu = 135, \sigma = 8)$. L'apartat a) és un càlcul de probabilitat. Recordar de l'assignatura anterior.

$$\begin{aligned} P(X \geq 145) &= P(\frac{X-135}{8} \geq \frac{145-135}{8}) \\ &= P(Z \geq 1{,}25) \\ &= 1 - \Phi(1{,}25) = 1 - 0{,}8944 = 0{,}1056 \end{aligned}$$

on Z denota la v.a. normal estàndar i Φ la funció de distribució de Z. L'apartat b) equival al càlcul del quantil 0.975 de X.

$$\begin{aligned} P(X \leq c) &= 0{,}975 \Rightarrow \\ P(\frac{X-135}{8} - \leq \frac{c-135}{8}) &= 0{,}975 \Rightarrow \\ P(Z \leq \frac{c-135}{8}) &= 0{,}975 \Rightarrow \\ \Phi(\frac{c-135}{8}) &= 0{,}975 \Rightarrow \\ \frac{c-135}{8} &= \Phi^{-1}(0{,}975) \end{aligned}$$

De la funció de distribució de Z resulta que

$$\Phi^{-1}(0{,}975) = 1{,}96$$

Per tant,

$$\frac{c-135}{8} = 1{,}96$$

I, finalment

$$c = 135 + 8 \cdot 1{,}96 = 150{,}68$$

Apartat c). Per resoldre aquest apartat cal tenir en compte les propietats de la mitjana mostral en poblacions normals. En general, si X és $N(\mu, \sigma)$ aleshores, en mostres de mida n, $\bar{X}$ és $N(\mu, \frac{\sigma}{\sqrt{n}})$. En el problema en qüestió, $n = 16$ i la mitjana mostral descriu l'alçada mitjana que observariem en mostres de 16 nens. Tenim , $\bar{X}$ és $N(135, 2)$. Per tant,

$$\begin{aligned} P(\bar{X} \geq 145) &= P(\frac{\bar{X}-135}{2} \geq \frac{145-135}{2}) \\ &= P(Z \geq 5) \approx 0 \end{aligned}$$

Observa que respecte l'apartat a), la probabilitat d'aquest succés ha disminuït, això és conseqüència de que X i $\bar{X}$ són normals amb la mateix mitjana però amb desviacions estàndar força diferents 8 i 2, respectivament. Apartat d). Hem de trobar el quantil 0.975 de $\bar{X}$.

$$\begin{aligned}
P(\bar{X} \leq c) &= 0{,}975 \Rightarrow \\
P(\frac{\bar{X} - 135}{2} - \leq \frac{c - 135}{2}) &= 0{,}975 \Rightarrow \\
P(Z \leq \frac{c - 135}{2}) &= 0{,}975 \Rightarrow \\
\Phi(\frac{c - 135}{2}) &= 0{,}975 \Rightarrow \\
\frac{c - 135}{2} &= \Phi^{-1}(0{,}975)
\end{aligned}$$

De la funció de distribució de Z resulta que

$$\Phi^{-1}(0{,}975) = 1{,}96$$

Per tant,

$$\frac{c - 135}{2} = 1{,}96$$

I, finalment

$$c = 135 + 2 \cdot 1{,}96 = 138{,}92$$

Observa que en comparació a l'apartat b) hem passat de $c = 150{,}68$ a $c = 138{,}92$. Això també reflecteix el fet que $\bar{X}$ en comparació amb X té una funció de densitat de probabilitat més "concentrada" en la mateixa mitjana.

Exercici 4

L'alçada de les plantes, X, segueix una distribució $N(\mu = 70, \sigma = 5)$, segons certifica el centre de jardinaria. En mostres de mida 20 de X, la distribució de

$$\frac{19S^2}{5^2}$$

on S^2 és la variancia mostral (corregida), és una khi-quadrat amb 19 graus de llibertat, χ^2_{19}. Ara, només ens cal calcular

$$\begin{aligned}
P(S^2 < 10) &= P(\frac{19S^2}{25} < \frac{190}{25}) \\
&= P(\chi^2_{19} < 7{,}6) = 0{,}01
\end{aligned}$$

Exercici 5

A la població de nens, estem estudiant una característica poblacional com és patir una malaltia congènita, aquesta malaltia té probabilitat de presentar-se $p = 1/400$. És a dir, aquesta és la probabilitat que un nen de la població d'estudi elegit a l'atzar presenti la malalatia.

Apartat a). En mostres de mida 800, la variable aleatòria X que compta el nombre de nens afectats té distribució Binomial($n = 800, p = 0{,}0025$). I té distribució aproximada Poissonl($\lambda = n \cdot p = 800 \cdot 0{,}0025 = 2$) . Ens serà convenient fer el càlcul amb la distribució aproximada

$$P(3 \leq X \leq 5) \approx P(X=3)+P(X=4)+P(X=5) = e^{-2}\frac{2^3}{3!}+e^{-2}\frac{2^4}{4!}+e^{-2}\frac{2^5}{5!} = 0{,}307$$

Apartat b). En mostres de mida 600, la variable aleatòria X que compta el nombre de nens afectats té distribució Binomial($n = 600, p = 0{,}0025$). I té distribució aproximada Poissonl($\lambda = n \cdot p = 600 \cdot 0{,}0025 = 1{,}5$) . Ens serà convenient fer el càlcul amb la distribució aproximada. Observa que es tracta de calcular una probabilitat condicionada

$$\begin{aligned} P(X=3|X \geq 2) &= \frac{P(\{X=3\} \cap \{X \geq 2\})}{P(X \geq 2)} \\ &= \frac{P(X=3)}{P(X \geq 2)} \\ &= \frac{e^{-1{,}5}\frac{(1{,}5)^3}{3!}}{1 - e^{-1{,}5}\frac{(1{,}5)^0}{0!} - e^{-1{,}5}\frac{(1{,}5)^1}{1!}} = 0{,}284 \end{aligned}$$

Apartat c). En mostres de mida 4000, la variable aleatòria X que compta el nombre de nens afectats té distribució Binomial($n = 4000, p = 0{,}0025$). Aquesta distribució queda aproximada per una Poissonl($\lambda = n \cdot p = 4000 \cdot 0{,}0025 = 10$) però també per una $N(\mu = \lambda = 10, \sigma = \sqrt{\lambda} = \sqrt{10})$.

$$\begin{aligned} P(X \geq 13) &\approx P(\frac{X-10}{\sqrt{10}} \geq \frac{13-10}{\sqrt{10}}) \\ &= P(Z \geq \frac{3}{\sqrt{10}}) = 0{,}1711 \end{aligned}$$

Exercici 6

Denotem amb X la v.a. que descriu el succés "el consumidor coneix el producte". Clar, X segueix una distribució Bernoulli de probabilitat $p = 1/4$, segons l'agència. Si considerem mostres de 1000 consumidors, podem definir la v.a. S, nombre de consumidors que coneixen el producte. S té distribució binomail de paràmetres $n = 1000$ i $p = 1/4$. Conseqüència del teorema central de límit, una distribució aproximada de S és la $N(\mu = np, \sigma = \sqrt{npq})$. Per tant, $\mu = 250$ i $\sigma = \frac{5}{2}\sqrt{30}$.

$$\begin{aligned} P(S \geq 232) &= P(\frac{S-250}{\frac{5}{2}\sqrt{30}} \geq \frac{232-250}{\frac{5}{2}\sqrt{30}}) \\ &= P(Z \geq -1{,}31) \\ &= P(Z \leq 1{,}31) = \Phi(1{,}31) = 0{,}905 \end{aligned}$$

Exercici 7

Apartat a) Si la variable aleatòria que descriu la distribució dels valors poblacionals és $N(\mu_X = 100, \sigma_X = 10)$. Aleshores en mostres de mida 25 la mitjana mostral $\bar{X}$ té distribució $N(\mu_{\bar{X}} = 100, \sigma_{\bar{X}} = \frac{10}{5} = 2)$. Calculem

$$P(\bar{X} > 95) = P(\frac{\bar{X}-100}{2} > \frac{95-100}{2}) = P(Z > -2{,}5) = 0{,}9938$$

Apartat b)

En mostres de mida n la mitjana mostral $\bar{X}$ tindrà distribució $N(\mu_{\bar{X}} = 100, \sigma_{\bar{X}} = \frac{10}{\sqrt{n}})$. Ens demanen que trobem n tal que

$$P(|\bar{X} - 100| < 5) \geq 0{,}95$$

De manera equivalent cal que es verifiqui

$$P(|\frac{\bar{X}-100}{10}\sqrt{n}| < \frac{5}{10}\sqrt{n}) \geq 0{,}95$$

I, per tant

$$\begin{aligned} P(|Z| < \frac{1}{2}\sqrt{n}) &= \geq 0{,}95 \Rightarrow \\ P(-\frac{1}{2}\sqrt{n} < Z < \frac{1}{2}\sqrt{n}) &= \geq 0{,}95 \Rightarrow \\ \Phi(\frac{1}{2}\sqrt{n}) - \Phi(-\frac{1}{2}\sqrt{n}) &= \geq 0{,}95 \Rightarrow \\ 2\Phi(\frac{1}{2}\sqrt{n}) - 1 &= \geq 0{,}95 \end{aligned}$$

És a dir, cerquem n tal que

$$\Phi(\frac{1}{2}\sqrt{n}) \geq 0{,}975 \Rightarrow \frac{1}{2}\sqrt{n} = \Phi^{-1}(0{,}975)$$

Per tant (veure taules de la $N(0,1)$), resulta

$$\frac{1}{2}\sqrt{n} \geq 1{,}96$$

Finalment,

$$n \geq (2 \cdot 1{,}96)^2 = 15{,}3664$$

És suficient prendre $n = 16$.

Exercici 8

Segons la companyia FLYFLY, si un passatger disposa de reserva, la probabilitat que es presenti per embarcar és $p = 0{,}06$. Si tenim 260 reserves, el nombre de passatgers que es presentaran per embarcar és una variable aleatòria X amb distribució Binomial de paràmetres $n = 260$ i $p = 0{,}06$. Podem aproximar la distribució de X d'acord a una normal de paràmetres $\mu = np = 260 \cdot 0{,}06 = 244{,}4$ i $\sigma = \sqrt{np(1-p)} = \sqrt{260 \cdot 0{,}06 \cdot 0{,}94} = 3{,}83$. La companyia podrà acomodar a tots els passatgers que es presentin quan $X \leq 250$, per tant

$$P(X \leq 250) = P(\frac{X - 244{,}4}{3{,}83} \leq \frac{250 - 244{,}4}{3{,}83}) = P(Z \leq 1{,}46) = \Phi(1{,}46) = 0{,}93.$$

Exercici 9

Apartat a) La mitjana mostral $\bar{X}_n$ té distribució $N(\mu, \frac{2}{\sqrt{n}})$, per tant puc escriure la desigualtat de l'enunciat de manera equivalent:

$$\begin{aligned}
P(|\frac{\bar{X} - \mu}{2}\sqrt{n}| \leq \frac{0{,}1}{2}\sqrt{n}) &\geq 0{,}95 \Rightarrow \\
P(|Z| \leq \frac{0{,}1}{2}\sqrt{n}) &\geq 0{,}95 \Rightarrow \\
P(-\frac{0{,}1}{2}\sqrt{n} \leq Z \leq \frac{0{,}1}{2}\sqrt{n}) &\geq 0{,}95 \Rightarrow \\
2\Phi(\frac{0{,}1}{2}\sqrt{n}) - 1 &\geq 0{,}95 \Rightarrow \\
\Phi(\frac{0{,}1}{2}\sqrt{n}) &\geq 0{,}975
\end{aligned}$$

Per tant, de les taules

$$\frac{0{,}1}{2}\sqrt{n} \geq 1{,}96 \Rightarrow n \geq (\frac{2 \cdot 1{,}96}{0{,}1})^2 = 1536{,}64$$

Hauriem de prendre n=1537.

Apartat b)

Cal tenir en compte que

$$E(|\ \bar{X}_n - \mu\ |^2) = \mathrm{var}(\bar{X}_n) = \frac{4}{n}$$

Aleshores,

$$\frac{4}{n} \leq 0{,}1 \Rightarrow\ n \geq \frac{4}{0{,}1} = 40$$