Alicia Morote
http://gemma.atipic.net

Analisi de Dades en Psicologia I
AC2

- **PRESENTACIÓ I CORRECCIÓ PAC 2**

PRESENTACIÓ I CORRECCIÓ PAC 2 :

Enunciat A

A la població la distribució de la variable "SATISFACCIO" obtingudes pels subjectes a les respostes dels qüestionaris segueix una llei normal de mitjana 6 punts i desviació de 2 punts.

Exercicis enunciat A:

1. **Quin percentatge de subjectes d'aquesta població té una puntuació de SATISFACCIO superior a 7 punts?**

   Hem de trobar l'àrea que hi ha per sobre del valor 7 estandarditzat. Per tant, primer estandarditzem el valor 7:

   $z_{SATISFACCIÓ=7} = (7-6)/2=0,5$

   A continuació trobem l'àrea que hi ha per sobre de z:

   $p(z \geq 0,5) = p(z \leq -0,5)=$ **0,3085**

   INTERPRETACIÓ

   El percentatge de subjectes d'aquesta població que obté puntuacions superiors a 7 punts en SATISFACCIO és de 30,85%.

   Una altra forma de calcular-ho és:

   $z=(7-6)/2=0.5$; $p(z \leq 0.5) = 0.6915$; $1 - 0.6915 =$ **0,3085**

   (Si busquem a la taula de la distribució normal estandarditzada trobem que el valor 0.5 dóna una probabilitat de 0.6915. Com que volem conèixer el percentatge que obté puntuacions superiors: 1 – 0,6915 = 0,3085).

2. **Quina és la probabilitat de que un subjecte d'aquesta població tingui una puntuació de SATISFACCIO entre 8 i 9 punts?**

   Estandarditzem el valor 8:

   $z_{SATISFACCIÓ=8} = (8-6)/2=1$; $p(z \leq 1) = 0.8413$

   Estandarditzem el valor 9:

   $z_{SATISFACCIÓ=9} = (9-6)/2=1.5$; $p(z \leq 1.5) = 0.9332$

   Per tant,

   $p(z \in [1,1.5] = p(z \leq 1.5) - p(z \leq 1) = 0.9332 - 0.8413 =$ **0,0919**

   INTERPRETACIÓ

   La probabilitat de que agafant un subjecte d'aquesta població a l'atzar tingui una puntuació de SATISFACCIÓ entre 8 i 9 punts és de un 9,2%.

3. **Quin valor deixa per sota el 95% de les puntuacions de SATISFACCIO de la població?**

   Observem en el llistat d'àrees que el valor que més s'acosta a **0.95 (95%) és 0.9505**. Per filar més prim faig el càlcul amb l'Excel amb la funció **DISTR.NORMAL.ESTAND.INV**. El resultat del valor z és de: **1,644853**

   0.9505 és l'àrea sota la corba d'un valor z inferior o igual a 1.64. Ara hem de desestandarditzar aquest número:

   $z=(x-\mu)/\sigma$; $1.64=(x-6)/2$; $x=1.64 \cdot 2+6=$ **9,28**

   INTERPRETACIÓ

   A la població trobem que el 95% de les puntuacions de SATISFACCIÓ està aproximadament per sota del 9,3

4. Entre quins valors de la variable SATISFACCIO es troba el 90% central dels individus d'aquesta població?

Si agafem el 90% central de valors, estem excloent un 5% per cada cua de la distribució **(5% · 2 = 10%).** Així doncs, ens interessa conèixer quina satisfacció correspon als valors estandarditzats z(p=0.050) i z(p=0.950).

$z_{0.050}$ = **–1.64**

x = –1.64 · 2 + 6 = 2,72

$z_{0.950}$ = **+1.64**

x = 1.64 · 2 + 6= 9,28

INTERPRETACIÓ

Per tant, el 90% central dels individus de la població tindran una puntuació de SATISFACCIÓ entre 2.72 i 9.28 punts.

A més, per a qualsevol corba de densitat normal si prenem 1,645 desviacions estàndard a cada banda de la mitjana aritmètica obtenim el 90% de l'àrea. Si fem els càlculs amb tots els decimals varia molt lleugerament el resultat:

1,645 · 2 = 3,29 6 – 3,29 = 2,71 6 + 3,29 = 9,29

En aquest cas el 90% central d'aquesta població es troba compresa entre els valors **2,71** i **9,29** de satisfacció.

5. Quins són els tres quartils de la distribució d'aquesta variable?. Entre quins valors de la variable es troba el 50% central dels individus d'aquesta població?

Començaré pel primer quartil que podem calcular-lo de vàries maneres:

a. El **primer quartil** és aquell valor que deixa el **25%** de la distribució per sota (Percentil 25). Per a calcular aquest valor estandarditzem el valor x restant la mitjana aritmètica i després dividint el resultat per la desviació estàndard.

Utilitzant la taula "**Àrees sota la corba normal estàndard**" trobem que el valor z al que li correspon un percentatge de 0.2514 sobre 1, és a dir, un 25.14% (el més proper a 25) és de **–0.67**. Per tant:

z= - 0,67 - 0,67=(x-6)/2 -1,34=x–6 x=6–1,34 x=4,66

b. Utilitzem l'Excel per a obtenir un resultat més exacte. Fem el següent:

z(p=0,25) = **- 0,674**

x= - 0,674·2+6=7,34 -1,348+6 = 4,652

Per tant, el primer quartil ens dóna el valor 4,65

El segon quartil coincideix amb la mediana, i en una corba de densitat normal la mediana i la mitjana aritmètica coincideixen, així doncs **el valor del segon quartil és 6**.

El tercer quartil també el calcularem de dues maneres:

a. El **tercer quartil** és aquell valor que deixa el **75%** de les observacions per sota (Percentil 75). Fem el mateix procés que per a calcular el primer quartil. Utilitzant la taula "**Àrees sota la corba normal estàndard**" trobem que el valor z al que li correspon un percentatge de 0,7486 sobre 1, és a dir, un 74,86% (el més proper a 75) és de **0,67**, que coincideix amb el valor del primer quartil però en positiu. Per tant:

z=0,67 0,67=(x-6)/2 1,34=x–6 x=6+1,34x=7,34

b. Ara ho comparem amb el resultat de l'Excel i dóna el mateix. Fem el següent:

x(p=0,75) = **0,674**

x = 0,674 · 2+6 = 7,348

En resum, els tres quartils són:

el primer 4,65, el segon 6 i el tercer 7,34

El 50% central es troba entre el primer i el tercer quartil, per tant entre els valors 4,65 i 7,34 .

6. **Si es considera que el 30% dels subjectes amb puntuació de SATISFACCIO més altes es poden considerar "molt satisfets", quina mitjana de SATISFACCIÓ determinarà el punt de tall que permet classificar als subjectes com a " molt satisfets "?**

Per tal de trobar el valor de subjectes "molt satisfets" per sobre del qual es troba el 30% de subjectes amb puntuacions més altes i per sota del qual es troba el 70% restant busquem a la taula "Àrees sota la corba normal estàndard" el valor z al que li correspon el percentatge més proper a 70%. Aquest valor és **0,52** i dóna un percentatge de **0,6985** sobre 1, és a dir, un **69,85%**. Per tant,

z=0,52 0,52=(x-6)/2 0,52 · 2= x - 6 x= 1,04+6 x= 7,04

També podem fer el càlcul amb l'Excel mitjançant la fórmula **DISTR.NORMAL.ESTAND.INV**. que ens dóna el mateix resultat:

0,524401003 $= z_{0.7}$

x=0,52 · 2 + 6 = 7,04

INTERPRETACIÓ

La mitjana de SATISFACCIÓ que determinarà el punt de tall, que permet classificar als subjectes com a "molt satisfets", es troba en 7,04 punts.

7. **Si agafem mostres de grandària 70. Entre quins valors de la variable mitjana de SATISFACCIO es situaran el 90% central de la distribució de les mitjanes observades en mostres de grandària 70, extretes de la població?**

El 90% central de valors exclou un 5% per cada cua de la distribució **(5% · 2 = 10%).** Així doncs, ens interessa conèixer quina satisfacció correspon als valors estandarditzats z(p=0.050) i z(p=0.950), que ja hem calculat a l'exercici 4.

D'altra banda, l'error estàndard és: $\sigma_{\bar{x}} = \frac{2}{\sqrt{70}} =$ **0,23904572**

$z_{0.050}$ **= -1,64**

x = -1,64 · 0,239 + 6 = 5,608

$z_{0.950}$ **= +1,64**

x = 1,64 · 0,239 + 6= 6,392

O bé, **6 ± 1,64 · 0,0239 = 5,608; 6,392**

INTERPRETACIÓ

El 90% central dels valors de la distribució de mitjanes de mostres de grandària 70, extretes d'aquesta població, es trobarà aproximadament entre 5,6 i 6,4 punts de SATISFACCIÓ.

8. **Si escollim mostres de 70 subjectes, quina serà la probabilitat que en escollir a l'atzar una d'aquestes mostres, la puntuació superi els 5 punts?**

Sabem, perquè l'hem calculat anteriorment, que **l'error estàndard** és **0.239**. (Veure exercici anterior). A continuació estandarditzem el valor 5:

z **= (x – μ) / σ;** ***z*** **= (5 – 6) / 0.239 = - 4,184**

A la taula de la "distribució normal estandarditzada" no trobem el valor ***z*** **= - 4.184**, però com que en els extrems la corba de distribucions és infinita obtindríem **- 0,0002**. Volem saber el percentatge que obté puntuacions superiors, per tant: **1 – 0,0002 = 0,9998**. És a dir:

p(*z* ≤ - 4.18) = 0,0002; p(*z* ≥ - 4,18) = 1 - 0,0002 = 0,9998

Per estar-ne més segura i comprovar si ens dóna el mateix, faig el càlcul amb l'Excel amb la funció **DISTR.NORM.ESTAND.** i el resultat és: **0,00001433**

Per tant, **1 – 0,00001433 = 0,9998**

INTERPRETACIÓ

La probabilitat que en escollir a l'atzar una d'aquestes mostres la mitjana de la puntuació en SATISFACCIÓ superi els 5 punts és de 0,9998 (99,98%).

Enunciat B

En una segona fase s'agafa una mostra de 150 subjectes. S'ha obtingut una puntuació de SATISFACCIO de 5,5 punts. Sabem que la desviació de la població (σ) és de 2 punts.

Exercicis enunciat B:

9. Estimar la mitjana poblacional amb un nivell de confiança del 95% (α = 5%), i amb un nivell de confiança del 99% (α = 1%). Canvien els límits de l'interval de confiança?. Per què?

N = 150 subjectes $z_{0.025} = -1{,}96$ $Z_{0.975} = 1{,}96$ $\sigma = 2$ $x_0 = 5{,}5$

$Z_{0.005} = -2{,}57$ $Z_{0.995} = 2{,}57$

Error Standard:

$\sigma_x = \sigma / \sqrt{n}$ $\sigma_x = 2/\sqrt{150}$ $\sigma_x = 2/12.25$ $\sigma_x = \mathbf{0{,}16}$

Interval de confiança del 95% de la mostra:

5%: 2,5% a cada extrem, busquem a la taula $Z_{0.025} = \mathbf{-1{,}96}$ i $Z_{0.975} = \mathbf{1{,}96}$

$5{,}5 + (1{,}96 \cdot 0.16) = 5.5 + 0.31 = \mathbf{5{,}81}$

$5{,}5 - (1{,}96 \cdot 0.16) = 5.5 - 0.31 = \mathbf{5{,}19}$

INTERPRETACIÓ

La mitjana poblacional amb un nivell de confiança del 95% oscil·la entre els valors 5,81 i 5,19 de SATISFACCIÓ.

Interval de confiança de 99% de la mostra:

1%: 0,5% a cada extrem, busquem a la taula $Z_{0.005} = \mathbf{-2{,}57}$ i $Z_{0.995} = \mathbf{2{,}57}$

$5{,}5 + (2{,}57 \cdot 0{,}16) = 5.5 + 0.41 = \mathbf{5{,}91}$

$5{,}5 - (2{,}57 \cdot 0{,}16) = 5.5 - 0.41 = \mathbf{5{,}09}$

INTERPRETACIÓ

La mitjana poblacional amb un nivell de confiança del 99% oscil·la entre els valors 5,91 i 5,09 de **SATISFACCIÓ.**

Els límits de l'interval de confiança són diferents en cada cas, ja que si augmentem en nivell de confiança (probabilitat de que la mitjana es trobi entre aquests valors) augmenta també l'interval de confiança (els límits entre els quals trobem la mitjana poblacional). Si volem més precisió cal ampliar els marges.

10. Quants subjectes hauria de tenir la mostra per a poder fer una estimació per interval, amb un nivell de confiança del 99% i una precisió de 0,5 punts?

Marge d'error = $z_{\alpha/2}(\sigma / \sqrt{n})$

$n = [z_{\alpha/2}\,(\sigma / \text{marge d'error})]^2$

$n = [2.57\,(2 / 0.5)]^2 = 105{,}678 \approx 106$

$0.5 = 2{,}57 \cdot (2/\sqrt{n})$

$\sqrt{n} = 2{,}57 \cdot 2 / 0{,}5$

$n = 2{,}57^2 \cdot 2^2 / 0{,}5^2$

$n = 6{,}60 \cdot 4 / 0{,}25$

n = 26,4 / 0,25

n = 105,6

INTERPRETACIÓ

Per tant, per a poder fer una estimació per interval, amb un grau de confiança del 99% i una precisió de 0,5 punts caldria una mostra de 106 subjectes.