Atenció: Contesta en el full de respostes les solucions finals, pero entrega tots els fulls de càlculs. Raona sempre les respostes

---

- (1, 70 %) Problema íntegre: Analitzem la producció en cadena de resistències electròniques:

  Per a cada peça s'arreplega, entre d'altres, la resistivitat real (en kΩ), la màquina que l'ha realitzada (A o B) i la temperatura ambient en el moment de la fabricació.

  Les màquines estan regulades per a fabricar peces amb exactament 10.00 kΩ, encara que diversos factors afecten al valor final en cada unitat. Es consideren **defectuoses** les peces amb una resistivitat que **s'allunye més de** 0.01 kΩ del valor teòric. Preguntes:

- (1.1, 15 %) A partir dels següents estadístics de dues mostres arreplegades (una de la màquina A i altra de la B, unitats en kΩ)

	n	mitjana	desv.tip.	$x_{\text{mín}}$	$x_{0.01}$	$x_{0.05}$	$x_{0.10}$
A	40	9.9995	0.007	9.983	9.98408	9.9871	9.99
B	30	10.0006	0.0091	9.982	9.98254	9.9897	9.991

	$x_{0.25}$	$x_{0.50}$	$x_{0.75}$	$x_{0.90}$	$x_{0.95}$	$x_{0.99}$	$x_{\text{màx}}$
A	9.996	10	10.005	10.008	10.0083	10.01192	10.013
B	9.994	9.999	10.0055	10.0152	10.0176	10.02184	10.024

contesta explicant quins estadístics uses i per què:

- (1.1.1, 5 %) Quina màquina fa una producció més prop del estàndar?

Sol.: La A, perque es desvia 0.0005, menys que la B, que es desvia 0.0006.

- (1.1.2, 5 %) Quina màquina fa una producció més regular?

Sol.: La A, perque té una desviació típica menor (no cal calcular el coeficient de variació, perque les mitjanes són quasi iguals).

- (1.1.3, 5 %) Dóna el rang de possibilitats per al percentatge d'unitats defectuoses que es troba en la **mostra B**. Basa't momés en les dades de la taula.

Sol.: Com que la tolerància és 0.01, les peces defectuoses són aquelles amb valor < 9.99 o > 10.01. Per tant, com que $x_{0.05} < 9.99 < x_{0.10}$, aleshores hi ha entre un 5 i un 10 % de peces massa petites, i com que $x_{0.75} < 10.01 < x_{0.90}$, aleshores hi ha entre un 25 i un 10 % de peces massa grans. En total, com a mínim hi ha un 15 i com a màxim un 35 % de peces defectuoses.

- (1.2, 15 %) A partir de la petita mostra conjunta de temperatura i resistivitat per a la **màquina A**:

Temp	21.6	27.5	24.8	24.8	25.4	27.6	21.5	28.4
Resist	9.996	10.002	10.009	9.999	9.990	10.000	10.004	10.008

- (1.2.1, 10 %) Es pot dir que hi ha una relació lineal forta entre la temperatura i la resistivitat? En què et bases?

Sol.: Es pot calcular $r = 0.213838677$ (o R^2), que és molt baix. Per tant la relació lineal NO és forta.

- (1.2.2, 5 %) Independentment de la resposta anterior, usa la recta de regressió, per a predir la resistivitat que s'obtindria amb unes temperatures ambient de 25 i 40°C, i valora la qualitat de cada predicció.

Sol.: Temp $= 25$, Resist $= 10.00089814$: la qualitat és baixa per raó del baix r.

Temp $= 40$, Resist $= 10.00853789$: la qualitat és baixa per raó del baix r, però encara que r fóra alta, tampoc tindria qualitat atés que les dades són temperatures al voltant del 25, per tant molt lluny del 40.

- (1.3, 15 %) Assumint que la resistivitat pot ser modelitzada per la distribució normal, i tenint en compte només les dades de l'apartat 1.1, què es pot decidir amb base estadística sobre el correcte funcionament de la **màquina B**? És a dir, podem decidir que la **màquina B** funciona de la manera en la que està ajustada (amb la mitjana teòrica de 10.00 kΩ) usant una confiança del 90 %?

  Nota: *Imprescindible mostrar els resultats dels càlculs que donen suport a la decisió.*

Sol.: IC per $\mu = [10.0006 \pm (t_{29})_{0.95} \frac{0.0091}{\sqrt{30}}] = [10.0006 \pm 1.699 \times 0.0016614] = [10.0006 \pm 0.00282276] = [9.997777; 10.003422]$, per tant accepte que la mitjana $\mu = 10.00$ és correcta per estar dins l'IC.

- (1.4, 15 %) Assumint que la resistivitat pot ser modelitzada per la distribució normal, calcula la probabilitat (o percentatge a llarg termini) d'unitats defectuoses que fabricaria la **màquina B**.

  Nota: *Usa com a mitjana μ el valor teòric o el valor mostral, segons la conclusió de l'apartat anterior (és a dir, usa la mitjana teòrica si has concluït abans que la màquina funcionava bé, o el valor mostral si has concluït que la màquina no funcionava segons la mitjana teòrica). Tranquil que no afectarà si has fet mal l'apartat anterior. Usa com a variància σ^2 la variància mostral en qualsevol cas.*

Sol.: $P(9.99 < X < 10.01) = F(10.01) - F(9.99) = F_Z(1.09) - F_Z(-1.09) = 2F_Z(1.09) - 1 = 2 \times 0.8621 - 1 = 0.7242$. Per tant volem calcular el contrari, $1 - 0.7242 = 0.2758$.

- (1.5, 10 %) Per a la **màquina B**, calcula el **benefici esperat** que es pot calcular per una unitat abans de ser fabricada, que no es coneix encara si serà o no defectuosa, sabent que, si és correcta, comporta un benefici net de 1.45€, i si és defectuosa, una pèrdua neta de 0.15€.

  Nota: *Si necessites algun valor que depén d'algun apartat que no has pogut fer, inventa't un valor raonable (demana al professor durant l'examen per a assegurar-te de que és raonable), i explica'l sobre el paper.*

Sol.: Siga Y = 'benefici'. Aleshores, per la probabilitat de l'apartat anterior, $Y = 1.45$€ amb probabilitat 0.7242 i $Y = -0.15$€ amb probabilitat 0.2758. Usant ara la fòrmula de l'esperança:

$E(Y) = 1.45 \times 0.7242 + (-0.15) \times 0.2758 = 1.00872$€ .

- (2, 30 %) Altres problemes:

- (2.1, 15 %) Es programa un applet que usa un navegador d'internet de la manera següent: es comença en una pàgina web seleccionada per l'usuari (pot ser `www.uji.es`). Aleshores, l'applet detecta els enllaços existents i tria ún d'ells a l'atzar per a saltar a la pròxima pàgina web. I així succesivament fins que arriba a una pàgina que no té hypereferències, on s'atura. Suposa, per un moment, que saps que el nombre mitjà d'enllaços que tenen les pàgines web de la *World Wide Web* és exactament 2.302.

  - (2.1.1, 10 %) Calcula la probabilitat de que, en arribar a una pàgina web qualsevol desconeguda, l'applet finalitze perque no hi ha més hypereferències. Usa el model de distribució més raonable de entre els disponibles, a partir de les dades oferides.

Sol.: El model raonable és el de Poisson, perque s'ens dóna una mitjana per pàgina web, i pot agafar qualsevol valor de comptar des del 0 cap amunt (no hi ha límit a priori). Si X = 'nombre d'hiperenllaços en una pàgina web desconeguda', aleshores, $P(X = 0) = \mathrm{e}^{-2.302} = 0.1$.

  - (2.1.2, 5 %) Calcula la probabilitat de que l'applet visite més de 10 pàgines web abans de parar.
    Nota: *Si necessites algun valor que depén d'algun apartat que no has pogut fer, inventa't un valor raonable (demana al professor durant l'examen per a assegurar-te de que és raonable), i explica'l sobre el paper.*

Sol.: El nombre de pàgines que visitarà dependrà de si troba o no troba hiperenllaços en les pàgines que visita. Cada pàgina visitada por tindre hiperenllaços o no (èxit o fracàs), i sabem de l'apartat anterior la probabilitat de l'èxit (no tindre hiperenllaços, 0.1, o tindre hiperenllaços, 0.9). Per tant es tracta del tipus binomial o binomial negativa. De qualsevol manera, la probabilitat de visitar més de 10 pàgines és la probabilitat de trobar-se 10 èxits en els 10 primers intents (èxit = tindre hiperenllaços). Per tant, si Y = 'nombre de pàgines amb hiperenllaços en visitar 10 pàgines web', aleshores $Y \sim \text{Bin}(n = 10, p = 0.9)$ $P(Y = 10) = 0.9^{10} = 0.34867844$.

- (2.2, 15 %) Tras una temporada de recopilació de dades, s'analitza les extraccions que es fan en un caixer electrònic, de manera que les quantitats dispensades tenen una mitjana de 310€ i desviació típica de 80€.

  - (2.2.1, 5 %) Podries calcular, exactament o de manera aproximada, quina és la probabilitat de que, el pròxim usuari del caixer, traga més de 300€? Dóna una xicoteta explicació (a més de fer el càlcul, cas de ser possible).

    Sol.: No, perque no sabem quin model de distribució segueix la variable $X =$ 'quantitat dispensada per usuari'.

  - (2.2.2, 10 %) Podries calcular, exactament o de manera aproximada, quina és la probabilitat de que, entre els pròxims 100 usuaris del caixer, traguen més de 30000€? Dóna una xicoteta explicació (a més de fer el càlcul, cas de ser possible).

    Sol.: Sí, perque la variable $X =$ 'quantitat dispensada per 100 usuaris' és la suma de 100 variables desconegudes, però el TLC ens permet suposar que X segueix aproximadament la llei normal de mitjana $\mu = 100 \times 310 = 31000$ i variància $\sigma^2 = 100 \times 80^2 = 640000$. Per tant,

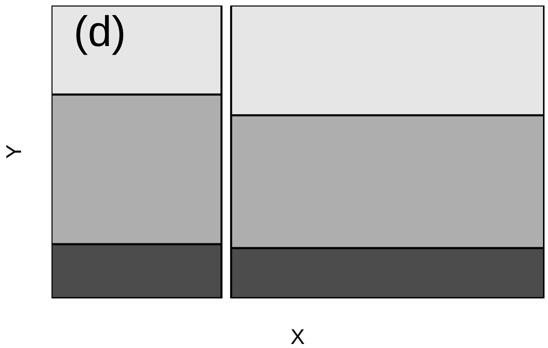
    $P(X > 30000) = 1 - F(30000) \approx 1 - F_Z(\frac{30000-31000}{800}) = 1 - F_Z(-1.25) = F_Z(1.25) = 0.8944$

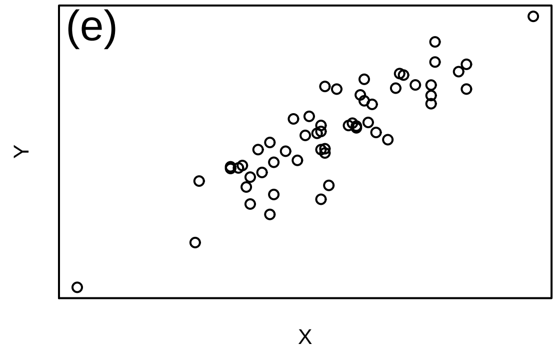
- (3, 10 %) Emparella cada lletra de frase amb una lletra de gràfic al qual fa referència, i escriu un petit raonament que explique la teua decisió (no puntúa sense el raonament correcte). Para atenció ja que està prou desordenat.

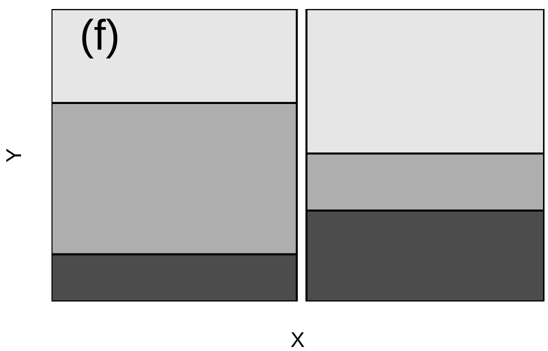
  (a) Les dades de la mostra A són més disperses que les de la mostra B

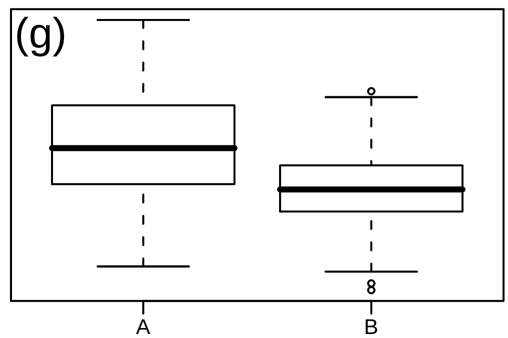
  (b) La relació lineal entre les variables X i Y és molt forta

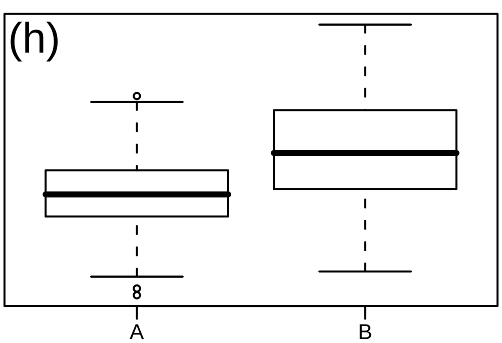
  (c) Les variables qualitatives X i Y són molt dependents

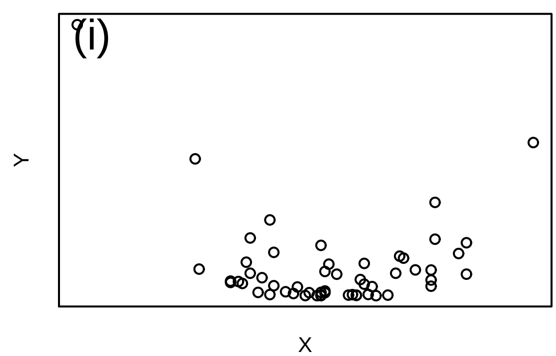












Sol.: (a)-(h), perque l'amplària de la caixa descriu la variabilitat.

(b)-(e), perque s'observa al núvol de punts una banda al voltant d'una recta.

(c)-(f), perque en cada barra del diagrama de barres s'observa una composició molt diferent, que indica la dependència d'una variable sobre l'altra.
