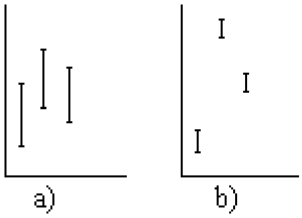
1.5. Estimar si un conjunt de mostres pertanyen a la mateixa població:

Objectius:

1. Obtenir un estimador inesbiaixat de la variança poblacional a partir de la mitjana de les variances d'un conjunt de mostres.
2. Obtenir un estimador inesbiaixat de la variança poblacional a partir de la variança de les mitjanes d'un conjunt de mostres pertanyents a la mateixa població.
3. Avaluar per anàlisi de variança si un conjunt de mostres independents pertanyen a la mateixa població.



Activitat 1.42. Si tenim un conjunt de mostres independents obtingudes per diferents procediments, en cas que aquests procediments siguen equivalents la dispersió entre les mostres haurà de ser proporcionada a la dispersió dins de cada mostra (figura a).Per tal de avaluar-ho, treballarem amb m mostres de grandària n i anomenarem:
X_{jk} a l'element k de la mostra j
$\overline{X}_j$ i $s(X_j)^2$ respectivament a la mitjana i la variança de la mostra j
$\overline{X}$a la mitjana de la mostra de grandària m·n resultant de mesclar les m mostres de grandària n.
Demostrar el
<u>Teorema 1.41:</u> $\overline{X}= \sum_j \overline{X}_j/m$.

Activitat 1.43. Anomenarem *variança dintre de variables* a la mitjana de les variances $s_w^2 = \sum_j s(X_j)2/m$.
Suposarem que totes les mostres pertanyen a poblacions si més no amb la mateixa variança σ^2 .
Demostrar el
<u>Teorema 1.42:</u> $\mu(s_w^2) = \sigma^2\cdot(n-1)/n$.
Anomenarem per tant *variança corregida dintre de variables* a $\hat{s}_w^2 = s_w^2\cdot n/(n-1)$, la qual serà un estimador inesbiaixat de la variança poblacional σ^2 .

Activitat 1.44. Tenint en compte el
<u>Teorema 1.43:</u> si les variables aleatòries independents Y_1, Y_2 tenen tenen distribució Xi-quadrat amb graus de llibertat ν_1 i ν_2 respectivament, aleshores la variable aleatòria Y_1+Y_2 té distribució Xi-quadrat amb $\nu_1+\nu_2$ graus de llibertat
demostrar el
<u>Teorema 1.44:</u> $mn\cdot s_w^2/\sigma^2$ té distribució Xi-quadrat amb m·(n-1) graus de llibertat.

Activitat 1.45. Anomenarem
$\mu_j=\mu(X_j)$ a la mitjana de la població a la qual pertany la mostra j
$\mu=\mu(X)$ a la mitjana de la població resultant de mesclar les poblacions a les quals pertanyen les m mostres
$\alpha_j=\mu_j-\mu$ para cada mostra j (naturalment, valdrà 0 si totes les mostres pertanyen a la mateixa població).

Demostrar que en qualsevol cas s'acompleix el Teorema 1.45: $\sum_j \alpha_j = 0$.

Activitat 1.46. Anomenarem *variança entre variables* a la variança de les mitjanes $s_b^2 = \sum_j (\overline{X}_j - \overline{X})^2/m = \sum_j \overline{X}_j^2/m - \overline{X}^2$.

Recordant de l'Activitat 1.26 que $\sigma(\overline{X}_j)^2 = \sigma^2/n$, $\sigma(\overline{X})^2 = \sigma^2/(mn)$, demostrar

Teorema 1.46: $\mu(\overline{X}_j^2) = \sigma^2/n + (\mu+\alpha_j)^2$

Teorema 1.47: $\mu(\overline{X}^2) = \sigma^2/(mn) + \mu^2$

Teorema 1.48: $\mu(s_b^2) = \sigma^2(m-1)/(mn) + \sum_j \alpha_j^2/m$.

Activitat 1.47. Anomenarem variança corregida entre variables a $\hat{s}_b^2 = s_b^2 \cdot nm/(m-1)$. Demostrar el

Teorema 1.49: $\mu(\hat{s}_b^2) = \sigma^2 + n \cdot \sum_j \alpha_j^2/(m-1)$.

Per tant, $\hat{s}_b^2$ serà un estimador inesbiaixat de la variança poblacional σ^2 si i sols si les poblacions a les quals pertanyen les diferents mostres tenen totes la mateixa mitjana (el que anomenem *hipòtesi nul·la* en la qual tot $\alpha_j=0$), cosa que naturalment passarà si totes les mostres pertanyen a la mateixa població. En aquest cas, $F=\hat{s}_b^2/\hat{s}_w^2$ haurà de ser proper a la unitat. En altre cas $\mu(\hat{s}_b^2)>\sigma^2$, i per tant es pot preveure que F siga major que la unitat.

Activitat 1.48. Tenint en compte que s_b^2 és la variança de la mostra ($\overline{X}_1$, $\overline{X}_2$,... $\overline{X}_m$), de grandària m, i que $\sigma(\overline{X}_j)^2 = \sigma^2/n$, demostrar que si les mostres pertanyen a la mateixa població s'acompleix el

Teorema 1.50: $mn \cdot s_b^2/\sigma^2$ té distribució Xi-quadrat amb m-1 graus de llibertat.

Activitat 1.49. Tenint el compte el

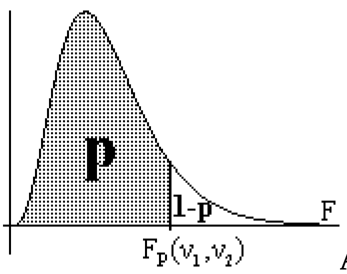
Teorema 1.51: si les variables aleatòries independents Y_1, Y_2 tenen tenen distribució xi-quadrat amb graus de llibertat ν_1 i ν_2 respectivament, aleshores la distribució de $F=(Y_1/\nu_1)/(Y_2/\nu_2)$ entre 0 i ∞ és

$W_{\nu1,\nu2}(F) = K_{\nu1,\nu2} \cdot F^{\nu_1/2-1}/(1+\nu_1 \cdot F/\nu_2)^{(\nu_1+\nu_2)/2}$, que s'anomena distribució *F de Snedecor* amb graus de llibertat ν_1 i ν_2 . $K_{\nu1,\nu2}$ s'escollís de manera que $\int_0^\infty W_{\nu1,\nu2}(F)\, dF = 1$

demostrar el

Teorema 1.52: si tenim m mostres independents de grandària n pertanyents a la mateixa població, aleshores

$F=\hat{s}_b^2/\hat{s}_w^2$ té una distribució F de Snedecor amb graus de llibertat $\nu_1=m-1$, $\nu_2=m \cdot (n-1)$.



Activitat 1.50. Si tenim m mostres independents de grandària n i

$F=\hat{s}_b^2/\hat{s}_w^2 > F_p(m-1, m \cdot (n-1))$, essent $F_p(\nu_1, \nu_2)$ el coeficient crític de la distribució F de Snedecor amb graus de llibertat ν_1 i ν_2 tal que la probabilitat d'un valor menor o igual a aquest coeficient siga p, aleshores podem rebutjar amb un nivell de significació $\beta=1-p$ la hipòtesis nul·la de que les mostres pertanguen a la mateixa població. Utilitzarem les tables de la distribució F de Snedecor (inversa) per determinar el corresponent

coeficient crític.
Problema 1.22: anotar el número de calcer en vàries mostres de la mateixa grandària entre l'alumnat assistent a classe i valorar si pertanyen a la mateixa població (a ser possible, procurar que alguna de les mostres estiga formada únicament per xics i altra únicament per xiques).

Activitat 1.51. Com hauríem d'interpretar el fet que $F=\hat{s}_b^2/\hat{s}_w^2 \ll 1$? Aplicar-ho a la resolució del següent
Problema 1.23: calcular l'estadístic F corresponent al següent parell de mostres:
X_1=(24'2, 25'3, 25'4, 26'2, 27'5)
X_2=(24'2, 25'3, 25'4, 26'2, 27'4)
Es pot considerar que les mostres no acomplisquen alguna de les premises del Teorema 1.52?
Per a valorar-ho amb un cert nivell de significació podem utilitzar la relació
$F_p(\nu_1, \nu_2)=1/F_{1-p}(\nu_2, \nu_1)$.

Treball 3 (per a la seua realització en equip)**:**
Comparar diferents procediments per a obtenir algun preparat químic utilitzant alguna variable aleatòria adequada (quantitat del preparat, temps per a la seua obtenció, etc.). Obtenir les dades d'experiències reals i utilitzar un mínim de 3 procediments aplicant com a mínim 5 vegades cada procediment. Estimar per anàlisi de variança si els diferents procediments es poden considerar equivalents.