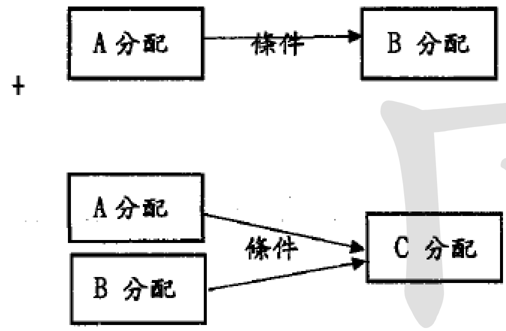
國立臺灣大學109學年度轉學生招生考試試題

題號： 23
科目：心理及教育統計學

題號： 23

1. 思考一下你所知道的<u>所有</u>隨機變數的分配 (distributions)，然後來回答下面的問題：
假設有 A 分配、B 分配，若 A 分配在某些條件之下會是 B 分配，請你用下面圖示的方式來表達出來，並請記得清楚寫出各分配的名稱及該些條件。有可能箭頭的左邊不只一個分配，也就是說可能有超過一個分配的組合 (如 A、B 分配)在某些條件下而變成另一個分配 (如 C 分配)。 [20%]



講義 P81~P82 分配間關係

※ F 分配的特性

1. 由於 F 為兩平方量之比值，故 $0 \leq F \leq \infty$，即 F 均為正值。
2. 每一對 df_1，df_2 就有一條分配曲線，如圖 9-5。
3. 與 χ^2 分配一樣，F 分配一為正偏分配，惟當 df_1，df_2 增加時，偏斜程度降低。
4. $df_2 \to \infty$時，F=1；當 $df_1 = 1$，$df_2 = \infty$時，$\sqrt{F}$ 會趨近於 Z 分配。
5. 曲線下面積為 1，故 F 分配為機率分配，其機率值可查 F 分配表。
6. F 分配常用來檢定兩常態母體之變異數是否相同。

$$\text{假設 } \sigma_1^2 = \sigma_2^2 \text{，則 } F = \frac{S_1^2}{S_2^2}$$

II. 各種分配的歸納與比較

(1) χ^2 分配、F 分配、t 分配的共同特性

i. 三者皆為小樣本分配。
ii. 三者皆為連續分配。
iii. 三者皆為最重要的抽樣分配。
iv. 三者所來自的母體皆為常態分配。
v. 三者的母體都是自由度運算而來。
vi. 三者皆能在某些條件下轉換成標準常態分配：

$df \to \infty,\ \sqrt{2\chi^2} \sim N.D.(\sqrt{2df-1}, 1)$
$df \to \infty,\ t \sim N.D.(0,1)$
$df_1 = 1,\ df_2 \to \infty,\ \sqrt{F} \sim N.D.(0,1)$

(2) t 分配與 Z 分配的比較

由 t 分配與標準常態分配(Z 分配)之重要表徵數的比較得知：

※ 相同處

1. 範圍在- ∞與+∞之間。
2. 以 0 為中心的左右對稱分配。

※ 相異處

1. t 分配的變異程度較大。
2. t 分配的峰態為高狹峰，Z 分配的峰態為常態峰。
3. 只有 t 分配會受自由度影響

(3) χ^2、F、t 統計量的定義與實用公式：

χ^2、F、t 統計量所來自之母體分配皆為常態分配，當母體之平均數 μ，變異數 σ^2 已知，則定義公式得以成立；但母數 μ、σ^2 常是統計推論的對象，此時就該使用實用公式。

(4) Z、χ^2、F、t 統計量間的關係

在自由度不同下，Z、χ^2、F、t 統計量間的關係為：

$$1.F_{1-\alpha(1,df)} = t^2_{1-\frac{\alpha}{2}(df)}$$

$$2.F_{1-\alpha(1,\infty)} = Z^2_{1-\frac{\alpha}{2}}$$

$$3.dfF_{1-\alpha(df,\infty)} = \chi^2_{1-\alpha(df)}$$

$$4.df/F_{1-\alpha(\infty,df)} = \chi^2_{\alpha(df)}$$

$$5.\chi^2_{1-\alpha(1)} = Z^2_{1-\frac{\alpha}{2}}$$

$$6.t_{1-\alpha(\infty)} = Z_{(1-\alpha)}$$

3. 我們會用樣本平均數（$\bar{X}$ ）來估計母群的平均數（μ），請從<u>五個角度</u>來討論樣本平均數（$\bar{X}$ ）是不是母群的平均數（μ）的一個好的估計值的？ [4% x 5=20%]

講義 P88~89 優良點估計值條件

二、良好點估計的判斷標準：

1. 不偏性(unbiasedness)

樣本統計量 $\hat{\theta}$的期望值等於母群參數 θ，即 E($\hat{\theta}$)=θ，則稱$\hat{\theta}$為 θ 的不偏估計值(unbiased estimator)。例如：E($\bar{X}$)=μ，

$E(\dfrac{\sum(X-\bar{X})^2}{n-1}) = \sigma^2, E(\dfrac{\sum(Y-\hat{Y})^2}{n-2}) = \sigma^2_{Y\cdot X}$, 所以$\bar{X}, \hat{S}^2, \hat{S}^2_{Y\cdot X}$分別是母群$\mu, \sigma^2, \sigma^2_{Y\cdot X}$的「不偏估計值」。

證明：$E(\hat{S}^2_X) = \sigma^2_X$

4. 甚麼是自由度(degree of freedom)? [5%] 就你所學過的統計（如 標準差、卡方檢定、t 檢定、迴歸、變異數分析等），在計算時怎麼定出自由度？（請把條件寫清楚） [15%]

講義 P77 自由度觀念

f. 自由度

自由度(degree of freedom,以 df 表示)(或 v 表示)
指樣本中能夠自由變動數值的個數。若樣本中有 n 個數值，則自由度 df=n，每增加一個限制(利用一個樣本估計值去推論一個母群參數)，則原來能自由變動數值就少掉一個，故自由度變為 df=n-1。

例: $(X_1+X_2+X_3+X_4+X_5)/n=\overline{X}$

$(X_1,\ldots\ldots X_5$，都可以自由指定，故 df=5)
當$(6+7+8+4+X_5)/5=6$

(當 $\overline{X}$ 指定為 6 時(因為母群平均數為 6)，任意指定 $X_1=6,X_2=7,X_3=8$；

$X_4=4$ 此時 X_5 就不能再自由指定了，而必須強迫為 5，因此自由度少掉一個，df=5-1)

5. 甚麼是中央極限定理 (Central Limit Theorem)? [10%] 它在推論統計的假設檢定(hypothesis testing)中扮演甚麼角色? [10%]

講義 P76 中央極限定理

d. 中央極限定理(Central limit theorem, CLT)以 $\overline{X}$ 分配為例
自平均數 μ，標準差為 σ 的母群中隨機抽取樣本大小為 n(n>30)的一組樣本

$(X_1,X_2,X_3,\ldots X_n)$，並求其平均數，稱為樣本平均數 $\overline{X}$，如此重複無限多次，即可得到無限多個 $\overline{X}$，則此無限多個 $\overline{X}$ 所形成的分配就稱為「$\overline{X}$ 分配」。
該分配為常態分配，期望值(或平均數)為 μ(E($\overline{X}$)=μ)，變異數為

$\frac{\sigma_x^2}{n}(V(\overline{X})=\frac{\sigma_x^2}{n})$，開根號為 $\frac{\sigma_x}{\sqrt{n}}$ 即標準誤。抽樣分配形成所依據的理論稱為「中央極限定理」。