

考試科目	統計學	所別	財政所	考試時間	2月23日(六) 第四節
------	-----	----	-----	------	--------------

一、假設有三個紙箱，紙箱一含 5 個紅球，6 個白球，紙箱二含 1 個紅球，2 個白球，紙箱三含 3 個紅球，2 個白球，茲隨機抽一個紙箱，再自其中隨機抽出一球，觀此球之顏色，試回答下列問題：

- (1) 若已知紙箱三被抽中，則紅球被抽中之條件機率為多少？(五分)
- (2) 求被抽中之球為紅色的機率。(五分)
- (3) 若被抽中之球為紅色，求其來自紙箱一之機率。(五分)
- (4) 試計算抽到紙箱一和抽到紅球這兩件事是否獨立？(五分)

二、 x 和 y 是兩個連續的隨機變數(continuous random variables)，他們的機率密度函數(probability density function)分別是 $f(x)$ 和 $f(y)$ ；他們的聯合機率密度函數(joint probability density function)則是 $f(x, y)$ ；他們的相關係數(coefficient of correlation)表示成 ρ_{xy} 。請分別證明下列的等式或者寫出其應有的式子：

- (1) $Var(x) = E(x^2) - E(x)^2$ (五分)
- (2) $Cov(x, y) = E(xy) - E(x)E(y)$ (五分)
- (3) 假設 a_1, a_2, b_1, b_2 都是固定的常數。請將 $Cov(a_1x + a_2y, b_1x + b_2y)$ 寫成 $Var(x), Var(y), Cov(x, y)$ 的函數型態。(五分)
- (4) ρ_{xy} 代表 x 和 y 的相關係數(coefficient of correlation)。假如 $y - E(y) = \beta[x - E(x)]$ ，則 x 和 y 的相關係數(coefficient of correlation)的絕對值等於 1。(五分)
- (5) 假若 $U = a_0 + a_1x, V = b_0 + b_1y, a_1b_1 > 0$ ，則 $\rho_{uv} = \rho_{xy}$ 。(五分)

三、假設樣本 $(y_1, \dots, y_n)$ 以及樣本 $(y_1, \dots, y_p)$ 都是從一個具有平均值(mean) μ 和變異數(variance) σ^2 的母體(population)隨機抽出的，其中 $p = n/2$ 。他們的樣本平均數(sample mean) $\bar{Y}_n$ 以及 $\bar{Y}_p$ 通常被視為母體平均數 μ 的可能估計值(estimator)；他們的樣本變異數(sample variance)則被用來當作母體變異數 σ^2 的可能估計值(estimator)。不過也有學者提出一個替代的樣本變異數 $\hat{\sigma}^2$ 來當作母體變異數 σ^2 的估計值(estimator)。這個替代的樣本變異數表示如下：

考試科目	統計學	所別	財政所	考試時間	2 月 23 日(六) 第四節
------	-----	----	-----	------	-----------------

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2}{n}$$

根據上述定義，請確認或修正下面的論述並證明之：

論述一： $\bar{Y}_n$ 與 $\bar{Y}_p$ 都是 μ 的不偏估計值(unbiased estimators)。(五分)

論述二： $\bar{Y}_n^2$ 是 μ^2 的一個不偏估計值(unbiased estimator)。(五分)

論述三：相對 $\bar{Y}_n$ 而言， $\bar{Y}_p$ 是一個有效率的估計值(efficient estimator)。(五分)

論述四： $\bar{Y}_n$ 不是 μ 的一致性估計值(consistent estimator)。(五分)

論述五：雖然 $\hat{\sigma}^2$ 是 σ^2 的一個偏估計值(biased estimator)，但卻是漸進的不偏估計值(asymptotically unbiased estimator)。(五分)

四、Mincer (1974) 使用 31,093 個美國白人男性去實證估計教育水準(years of schooling) X 對薪資(earning) Y 的影響。他主要想檢驗教育(education)是否存在正的經濟報酬(positive economic return)。他的線性迴歸模型(linear regression model)表示如下：

$$\log Y_i = \alpha + \beta X_i + e_i$$

其中被解釋變數(dependent variable) Y 是以對數型式(logarithmic form)表示。他得到的估計係數(estimated coefficients)以及估計係數的標準誤(estimated standard error)整理於下表：

	α	β
估計係數	$\hat{\alpha} = 7.58$	$\hat{\beta} = 0.070$
估計標準誤	$s_{\hat{\alpha}} = 3.20$	$s_{\hat{\beta}} = 0.00160$
估計係數的變異數	$Var(\hat{\alpha}) = A$	$Var(\hat{\beta}) = B$
假設檢定所需估計的 t 值	$t_{\hat{\alpha}} = C$	$t_{\hat{\beta}} = D$

備

註 試題隨卷繳交

考試科目	統計學	系別	財政所	考試時間	2 月 23 日(六) 第 ④ 節
------	-----	----	-----	------	-------------------

由於他的估計是根據一個非常大樣本($n=31,093$), 因此 t 分配的表訂數值(the tabulated values of the t distribution)跟標準常態分配(the standard normal)的表訂數值是一樣的。舉例來說, t 分配的表訂數值是 $t_{\alpha/2} = 1.96$, $\alpha/2 = 0.025$ 。請根據上述的訊息回答下列問題:

- (1)請計算表中的 A, B, C, D。(五分)
- (2)請建立 β 的 95%的信賴區間(confidence interval)。(五分)
- (3)請依他的研究目的幫忙建立教育經濟報酬的 null hypothesis 以及 alternative hypothesis。(五分)
- (4)請問這個研究主題的假說檢定(hypothesis testing)是屬於 one-sided test 或 two-sided test 呢? 並請根據各種數據說明他的估計結果在 5%的顯著水準下是支持或者拒絕(3)的 null hypothesis 呢。(五分)
- (5)你會如何解釋 $\hat{\alpha} = 7.58$ 以及 $\hat{\beta} = 0.070$ 的經濟涵義呢?(五分)
- (6)如果作者希望能夠直接從模型裡估計出薪資的教育彈性水準(education elasticity of earning), 他應該怎麼修改模型設定呢? 如果作者希望能夠直接從模型裡估計出教育對薪資的邊際效果(marginal effect), 他又應該怎麼修改模型設定呢?(五分)

備

註

試題隨卷繳交