八十九學年度統計所碩士班研究生招生考試

科目 統計學 科號 0303 共 4 頁第 1 頁 *請在試卷【答案卷】內作答

[redacted]: 所有考題均係填充題，只須將答案清楚列出（不要列計算過程），每一題考題空格中列有題號及配分，例如 (2a, 2 分) 表示題號第 2a 題，配分 2%。

在答案卷中必須清楚標明題號，無清楚題號者，一律不計分。

1. 設 $X_1,\ldots,X_n$ 爲一組從 Poisson (λ), $\lambda>0$ 取出之隨機樣本，則參數 $\theta=\ell n\sqrt{\lambda^3}$ 之最大概似估計量 (maximum likelihood estimator) 爲 ___(1a, 2 分)___，又 $(n-10^9)(\bar{X}-\lambda)/e^{100}\sqrt{(n+e^{100})^2\bar{X}}$ 之極限分布(limiting distribution) 爲 ___(1b, 3 分)___。

2. 假設 $X_1,\ldots,X_n$ 是一組從累積分佈函數 $F(x), -\infty<x<\infty$ 所取出的隨機樣本，定義 $Y=$ number of $X_1,\ldots,X_n\le C$，其中 C 爲一常數，則 $E(\frac{Y}{n})=$ ___(2a, 2 分)___，$Var(\frac{Y}{n})=$ ___(2b, 3 分)___。

3. 若 $Y_1<Y_2<Y_3$ 爲 X_1,X_2,X_3 之有序統計量(order statistics),其中 X_1,X_2,X_3 爲從機率密度函數(probability density function) $f(x)=2x,\ 0<x<1$ 所取出之一組隨機樣本。若 $Z_1=\frac{Y_1}{Y_2}$，$Z_2=\frac{Y_2}{Y_3}$，$Z_3=Y_3$，則 $E(Z_1Z_2Z_3)=$ ___(3, 5 分)___。

4. 若 f_0 代表 均勻 (uniform) $U[0,1]$ 之機率密度函數(probability density function)而 $f_1=4x$ 當 $0\le x\le\frac{1}{2}$，$f_1=4-4x$ 當 $\frac{1}{2}\le x\le 1$。現隨機變數 X 之機率密度函數可能爲 f_0 或 f_1，若依據此"單一"隨機變數來檢定 $H_0: f=f_0$ 對 $H_1: f=f_1$，則最佳棄卻域(best critical region)爲 ___(4, 5 分)___。設此檢定之顯著水準爲 α。

5. 假設 $X_1,\ldots,X_n$ 爲一組從常態 $N(\mu,\sigma^2)$ 所取出的隨機樣本，其中 μ 爲已知，而 σ^2 爲未知，則 σ 之最小充分統計量(minimum sufficient statistic)爲 ＿(5a, 2分)＿，又 σ 之 Rao-Cramer 下界(lower bound) 爲 ＿(5b, 3分)＿。

6. 假設 $X_1,\ldots,X_n$ 爲一組從 Bernoulli (θ) 所取出之隨機樣本，其中 θ 之先驗分布(prior distribution) 爲 $U[0,1]$。若損失函數(loss function)爲 $L(\hat{\theta},\theta)=\theta(\hat{\theta}-\theta)$，則 θ 之貝氏估計量(Bayes estimator) 爲 ＿(6, 5分)＿。

7. 若 $X_1,\ldots,X_n$ 爲一組從 Gamma (α,β) 所取出之隨機樣本，其中 $\alpha=3, \beta>0$，則 β 的一個等尾(equal-tail) $100(1-\alpha)\%$ 信賴區間爲 ＿(7, 5分)＿。（提示: 先求出 $\sum_{i=1}^{n} X_i/\beta$ 之分布）

8. 設 $X_1,\ldots,X_n$ 爲一組從機率密度函數(probability density function) $f(x;\theta)=\theta^p x^{p-1}e^{-\theta x}/\Gamma(p)$ 所取出之隨機樣本，其中 p 爲一已知正整數，θ 爲未知參數，則檢定 $H_0:\theta=\theta_0$ 對 $H_1:\theta>\theta_0, \theta_0$ 爲一常數之檢定力函數(power function) 爲 ＿(8, 5分)＿; 設此檢定之顯著水準爲 α。

9. 重複執行 Bernoulli 試驗一直到 k（已知）次成功發生，其中每次試驗之成功機率爲 θ，$0<\theta<1$ 且 θ 爲未知參數。若隨機變數 Y 表示總共需要的試驗次數，則 θ 之最小變異不偏估計量(minimum variance unbiased estimator) 爲 ＿(9, 5分)＿。（提示: 計算 $E\dfrac{k-1}{Y-1}$）

10.令隨機向量$(X,Y)'$爲具有二維常態(bivariate normal) $BN\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)$分布之隨機向量，其中$-\infty<\mu_1,\mu_2<\infty$，$0<\sigma_1^2$，$\sigma_2^2<\infty$，$0<\rho<1$，則隨機變數$X+Y$和$X-Y$爲相互獨立之充分必要條件爲 (10, 5分) 。

11.在民意測驗報告中針對候選人支持率之調查，經常有下列語句 "在信心水準爲95%之下，抽樣誤差爲 3%"，此信心水準 95%的意義爲何？（簡潔說明）(11a, 5分) 。假設某地區某候選人真正支持率爲p，今隨機歸還抽樣了n個人，其中有x人支持，則樣本支持率x/n之理論標準差 在二項分布成立時以n和p之函數表示爲 (11b, 3分) 。當$n=1600$, $x=160$ 時，樣本支持率的標準差估計值爲多少？ (11c, 2分，以百分數表之至小數二位)。若在作歸還抽樣前不知其支持率大約爲何，但又希望最大抽樣誤差不超過 2%，則至少須要抽多少人？(11d, 5分) 。

12.某地區過去一直男女人數差不多，但最近假設我們認爲男性可能已比女性人數爲多，想以數據加以檢定，設P爲當地男性之比率，則檢定之虛無假設(null hypothesis)爲 H_0: (12a, 3分)， 對立假設(alternative hypothesis)爲H_1: (12b, 3分) 。現任抽了 5 人，其中 3 人爲男性，則檢定之P-值(P-value)爲 (12c, 3分) 。若顯著水準設定爲 5%，則以上檢定之結論爲 (12d, 3 分) 。若任抽 10人，則必須至少 (12e, 3分) 人爲男性，才有充分之證據說明男多於女？

(可參考下值: 1/1024 = 0.00098, 45/1024=0.0439, 120/1024=0.1172)

13.在醫學的報導中，經常有類似下列的語句「根據研究的結果顯示 O 型者比非 O 型者易致胃潰瘍，其對比 (odds ratio, 又稱勝算比)估計值爲 2.5」。假設此組資料爲n_1個 O 型中有x_1個有胃潰瘍，n_2個非 O 型中有x_2個胃潰瘍，則對比估計值以n_1,n_2,x_1,x_2之函數表示爲＿(13a, 3 分)＿。現某人在另一地重複此類實驗，抽了m_1個 O 型且此m_1個人中有k_1個有胃潰瘍，並抽了m_2個非 O 型，此m_2個人中有k_2個有胃潰瘍，設對比之估計值爲$\hat{\alpha}$，則對比估計值之對數值，即 $\ln(\hat{\alpha})$，超過何值(以m_1,m_2,k_1,k_2之函數表示) (13b, 5 分)，可下結論在另一地之 O 型者比非 O 型者易致胃潰瘍。（假設顯著水準爲 5%，答案中可用適當分布之百分位點）。你需要的假設是 (13c, 5 分) 。

14.在上題某人重複之實驗中，設m_1個人血壓資料爲$(X_1,X_2,\ldots X_{m_1})$，其平均血壓爲$\bar{X}$，m_2個非 O 型者血壓資料爲$(Y_1,Y_2,\ldots Y_{m_2})$，其平均血壓爲$\bar{Y}$。若假設二樣本爲獨立，二母体血壓分布是常態且有相同之母体變異，則$\bar{X}-\bar{Y}$超過何值＿(14, 7 分)＿才可下結論 O 型母体平均血壓較非 O 型母体平均血壓爲高。（假設顯著水準爲 5%，答案中可用適當分布之百分位點）。