東吳大學 106 學年度碩士班研究生招生考試試題

系級	巨量資料管理學院碩士學位學程	考試時間	100 分鐘
科目	統計學	本科總分	100 分

1. (14 分) Let the random variable *X* has a Uniform(0,2) distribution, Y has a Uniform(0,1) distribution and the random variable Z has a Uniform(0, 0.5) distribution.

(a) 寫出並劃出 the probability density functions of *X*, *Y* and Z , $f(x)$, $g(y)$, and $h(z)$ with space (畫在同一個圖上，並標出f, g和 *h* 三個函數).

(b) 推導並劃出 and draw the cumulative distribution functions of *X* and *Y* and Z, i.e. $F(x)$, $G(y)$ and $H(z)$ (畫圖並標出F, G 和H三個函數).

2. (14 分)

(a) 請用式子或口頭說明中央極限定理 (Central Limit Theory)的內涵及其用途

(b) Let $\overline{X}$ be the mean of a random sample of $n = 25$ currents (in milliamperes) in a strip of wire in which each measurement has a mean of 15 and variance of 4. Calculate the approximate probability $P(\overline{X} < 15.784)$=?.

3. (14 分) Let $X_1, X_2, \cdots, X_n$ denote a series of random variables from a distribution with unknown parameter θ. 請說明概似函數 (likelihood function) 與聯合機率密度函數 (joint probability density function) 之相同處與相異處.

4. (14 分) 在進行統計假設檢定(hypotheses testing)以決定是否拒絕虛無假設(null hypothesis)時，經常會使用到 α (the significance level or tolerance level) 和 p-value 兩個數字之間的比較，請用文字敘述說明或公式或圖形，說明這兩個數字的意涵及其比較的背後邏輯

5. (14 分) Let *X* be a random variable associated with a Bernoulli trial, i.e. *X* has a Bernoulli distribution with the probability of success p .

(a) (4 分) Write down the probability mass function of *X* .

(b) (5 分) Derive the expected value of *X* .

(c) (5 分) Derive the variance of X .

6. (14 分) Prove the Theorem: If $X \sim N(\mu, \sigma^2)$, then $Z = (X - \mu)/\sigma$ is $N(0,1)$.

7. (16 分) Assuming that X_1 is $N(\mu_1, \sigma_1^2)$ with n_1 sample and X_2 is $N(\mu_2, \sigma_2^2)$ with n_2 sample. 寫出下列各虛無假設的檢定統計式及其所用以判斷檢定結果的分配

(a) $H_0: \mu_1 = 60$ (when σ_1^2 is unknown)

(b) $H_0: \mu_1 = \mu_2$ (when X_1 and X_2 are independent)

(c) $H_0: \mu_1 = \mu_2$ (when X_1 and X_2 are paired data taken from same subject and hence $n_1 = n_2 = n$)

(d) $H_0: \sigma_1^2 = \sigma_2^2$