

HW4

B10209040 陳彥倫

1.a.

信賴區間是統計學中一個重要的概念，用於估計母體參數（例如平均值或比例）的不確定性。信賴區間提供了一個範圍，我們可以合理地認為包含了真實母體參數值的範圍。具體地，信賴區間由兩個數值組成，分別稱為下限和上限。這兩個數值通常是根據樣本統計量計算出來的，例如樣本平均值或樣本比例。信賴區間的計算通常基於樣本的分佈，以及對母體參數的估計和不確定性的理解。信賴區間的定義可以用以下方式描述：「如果我們重複地從同一個母體中抽取多個樣本，然後對每個樣本計算一個估計值（如樣本平均值），然後按照某種方法對這些估計值進行分析，例如計算它們的平均值和標準差，然後根據這些統計量，我們可以構建一個區間。這個區間被稱為信賴區間，表示我們對真實母體參數值的估計具有一定的信心水平。」信賴區間通常以一個信心水平（confidence level）來描述，例如95%信心水平的信賴區間。這意味著在許多次抽樣中，我們可以期望有95%的信心，真實母體參數值位於所計算的信賴區間中。常見的信心水平包括90%、95%和99%等。信賴區間的計算方法取決於所研究的問題和樣本統計量。對於平均值的信賴區間，通常使用t分佈或正態分佈進行計算。對於比例的信賴區間，則使用二項分佈或正態分佈。

remark:

前半部的回覆概括了信賴區間在統計學上的概念及名詞解釋，而後提到信賴區間的定義通常以信賴水準來描述，與課堂上的理解相符。但對於t分布的使用並沒有詳細說明其通常是被用於母體變異數未知的情況。

1.b.

不合理，信賴區間之相關定義不適合用機率的概念描述，母體參數的真實數值落在信賴區間內的結果只有0跟1的區別。應該改為在100次的抽樣統計當中，有95次此區間涵蓋了某人在這次大選的真實支持率。

2.

此案例樣本數較大，且母體變異數未知，信賴區間的計算可套用：

$$\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}}$$

而題幹所代入之數字符合上述公式，此敘述正確。又因其符合95%信賴區間之定義，可以推斷[2.96,3.24]有0.95的機率包含真實平均。

3.

因為母體平均數即為樣本分配的平均，計算可得此區間約為[235.26,244.74]。而236位於此區間內，因此機器運作良好。

4.

此文章講述了威廉·希力·戈斯特（William Sealy Gosset）先生如何運用科學方法改進了啤酒釀製過程及其品質。為了得知瓶中酵母菌的確切數目，戈斯特採用統計學中的卜瓦松分布來描述，此分布的特性是平均數和變異數相同。。此分布適用於單位體積或時間內，稀有事件發生次數。此外卜瓦松發明的學生氏t分布也被用於小樣本的統計檢定，可以判斷樣本平均值是否與母體參數存在顯著差異。