補充題庫解說

主題五、主題六

（CH5～CH8）

Discrete Probability Distributions

離散型機率分配

常見離散型機率分配

分配類型	適用情境說明	分配函數
均勻分配 discrete uniform distribution	無論隨機變數如何變化，其機率值為固定常數，例如投擲一次骰子，機率都是1/6	$f(x) = \begin{cases} \frac{1}{n}, x = 1,2,..,n \\ 0, \text{other} \end{cases}$
二項分配 binomial distribution	在成功與失敗的機率不變情況下，不斷重複試驗所得的機率分配函數	$f(x) = C_x^n\, p^x\, q^{n-x}$, x=0,1,..,n 其中p為成功機率；q為失敗機率
伯努利分配 bernoulli distribution	在二項分配下，不論成功或失敗，只做一次（n=1），算是二項分配的特例	$f(x) = p^x\, q^{1-x}$, x=0,1 其中p為成功機率；q為失敗機率
超幾何分配 hyper geometric distribution	從N物中採抽出不放回抽樣，抽出n個樣本，其中N物中，有S個相同，N-S個不同。所以每次抽樣都會受前次抽樣影響	$f(x) = \frac{C_x^S\, C_{n-x}^{N-S}}{C_n^N}$, $0 \le x \le \min(n,S)$
波松分配 Poisson distribution	在一個單位時段或區段內，某事件發生次數（X）的問題	$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0,1,2 \ldots$ λ 為平均次數

其他離散型機率分配

分配類型	適用情境說明	分配函數
多項分配 multinomial distribution	二項分配的延伸，二項分配是把母體分成二種類別，而多項分配則將母體分類超過兩個以上	$f(x_1, x_2, \dots, x_k)$ $= \dfrac{n!}{x_1!\ x_2!\ \cdots\ x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$ $= \mathrm{C}^{\mathrm{n}}_{x_1,x_2,\dots,x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$
負二項分配 negative binomial distribution	第 r 次成功所需試驗的次數 （每次都是伯努利試驗） （X表示第 r 次事件發生時所需試驗次數）	$f(x) = C^{\mathrm{x}-1}_{\mathrm{r}-1}\ \mathrm{p}^{\mathrm{r}}\ \mathrm{q}^{\mathrm{x}-\mathrm{r}}$, $x = r, r+1, r+2, \dots..$ 其中p為成功機率；q為失敗機率
幾何分配 geometric distribution	不斷進行伯努利分配，直到第1次成功為止所需試驗次數之機率函數	$f(x) = \mathrm{q}^{\mathrm{x}-1}\mathrm{p}$, x=0,1,2,.... 其中p為成功機率；q為失敗機率

案例1

補5.3【台大財金】

三個死黨聚餐，以擲銅板決定付帳者，規定三者中不同者付帳，三者相同就再擲，

則第一次擲即成功的機率為 (1) ；

至少需擲三次才能作決定的機率為 (2)；

平均而言，需擲 (3) 次才能決定付帳者。

案例1解說

二項分配

(1) 令X表三人各自丟擲後出現正面的個數 → $X \sim B\left(\mathrm{n} = 3\,, \mathrm{p} = \frac{1}{2}\right)$

$$\mathrm{P}(\mathrm{X} = 1 \text{ or } \mathrm{X} = 2) = \mathrm{P}(\mathrm{X} = 1) + \mathrm{P}(\mathrm{X} = 2)$$

$$= \mathrm{C}_1^3 \left(\frac{1}{2}\right)^3 + \mathrm{C}_2^3 \left(\frac{1}{2}\right)^3 = \frac{3}{4}$$

幾何分配

(2) 令Y表直到決定付帳者為止所需丟擲次數 → $\mathrm{Y} \sim \mathrm{Geo}\left(\mathrm{p} = \frac{3}{4}\right)$

$$\mathrm{P}(\mathrm{Y} \geq 3) = 1 - \mathrm{P}(\mathrm{Y} = 1) - \mathrm{P}(\mathrm{Y} = 2) = 1 - \frac{3}{4} - \frac{3}{4} \times \frac{1}{4} = \frac{1}{16}$$

(3) 平均需擲幾次才能決定付帳者 → 求 Y 的期望值

$$E(\mathrm{Y}) = \frac{1}{\mathrm{p}} = \frac{1}{\left(\frac{3}{4}\right)} = \frac{4}{3} \quad (\text{次})$$

案例2

補5.7《成大財金》

一個在袋中有20個球，其中黑球有5個，白色球有15個，某統計教師確定以10次隨機抽取，抽中黑球數作為本學期「當」掉幾個學生的參考。若每次只抽一個球，令X代表抽中是黑球數，則在抽出放回情況下，X之機率函數為何？該教師最有可能「當」幾個學生？

案例2解說

(1)隨機變數X之機率分配 → $X \sim B\left(\mathrm{n}=10\,,\mathrm{p}=\frac{1}{4}\right)$

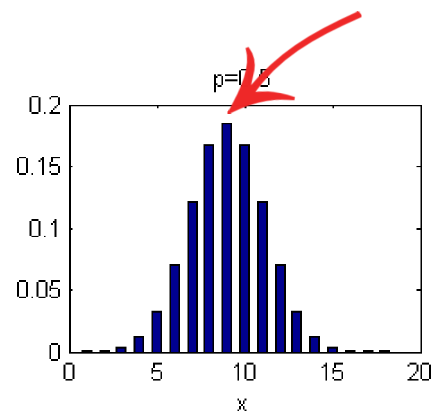
$$\therefore f(X)=C_x^{10}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{10-x}$$ ，x=0,1,2,3,.....,10

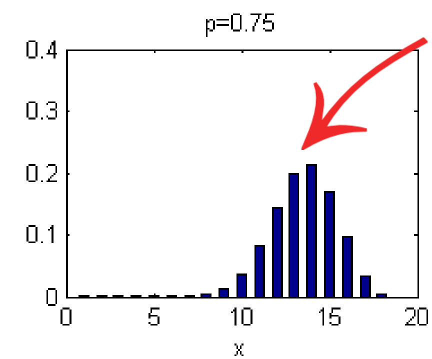
(2) 〝最有可能「當」掉幾個〞學生 → 求取「眾數」

二項分配的「眾數」：

$$M_0=\begin{cases}(\mathrm{n}+1)\mathrm{p} & \text{若}(\mathrm{n}+1)\mathrm{p}\text{ 不為整數}\\ (\mathrm{n}+1)\mathrm{p}-1\text{及}(\mathrm{n}+1)\mathrm{p}\,,\text{若}(\mathrm{n}+1)\mathrm{p}\text{為整數}\end{cases}$$

∴ 眾數M_0=(n+1)p
=(10+1)×0.25
=2.75 (不為整數)
M_0取2
→最有可能當掉2學生





Continuous Probability Distributions

連續型機率分配

常見連續型機率分配

分配類型	適用情境說明	分配函數
連續均勻分配 uniform distribution	隨機變數（X）在某連續區間內所發生的機率都相同	$f(x)=\frac{1}{b-a}$，$a \leq x \leq b$
常態分配 normal distribution	存在於大自然間的各種現象或狀態，都可以將母體視為常態分配	$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$ $-\infty < x < \infty$, μ為平均數, σ為標準差
標準常態分配 standard normal distribution	透過 $z=\frac{x-\mu}{\sigma}$ 的「標準化」，將原本要利用微積分計算求值，轉換成可以利用查表得到結果	$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$, $-\infty < z < \infty$
指數分配 exponential distribution	與Poisson隨機變數相反，指數分配的隨機變數（X）是描述連續兩事件發生的間隔時間	$f(x)=\frac{1}{\beta}e^{-\frac{x}{\beta}}$, $x \geq 0$ x 為第一次發生事件所需時間； β 為事件發生的平均時間；

其他連續型機率分配

分配類型	適用情境說明	分配函數
Gamma 分配 Gamma distribution	隨機變數（*X*）表示事件第 *a* 次發生所需的時間。因此若「*X>t*」表示事件第 *a* 次發生至少需要 *t* 個時間單位；另個說法是，在 *t* 時間內事件至少發生 *(a-1)* 次的機率	$f(x) = \frac{1}{\beta\Gamma(a)}(\frac{x}{\beta})^{a-1}\, e^{-\frac{x}{\beta}}$, $0 < x < \infty, a > 0, \beta > 0$ x 為第 *a* 次發生所需時間； β 表發生一次所需的時間；
卡方分配 chi-square distribution	卡方分配可以算是Gamma分配的特例（$a = \frac{\nu}{2}$，$\beta = 2$），在統計應用上，可進行單一母體變異數 σ^2 的統計推論；可用來做適合度檢定（goodness-of-fit test）、獨立性檢定（test of independence）與變異數齊一性檢定（test of homogeneity）	$f(x) = \frac{1}{2\Gamma\left(\frac{\nu}{2}\right)}(\frac{x}{2})^{\frac{\nu}{2}-1}\, e^{-\frac{x}{2}}$, $0 < x$，ν 為卡方分配的自由度

案例3

補6.4【成大國企】如果某次考試有4000人參加，成績X成常態分配，已知母體第2及第3四分位Q2＝63.25，Q3＝72.31，求：

(1)母體平均數 μ ＝？

(2)標準差 σ＝ ？

(3)第1四分位Q1＝ ？

(4)約有多少學生成績高於80分？

案例3解說

(1) $X \sim N(\mu, \sigma^2) \rightarrow \mu = Q_2 = 63.25$

(2) $P(X > Q_3) = P(X > 72.31) = P\left(Z > \frac{72.31 - 63.25}{\sigma}\right)$

$\rightarrow \frac{72.31 - 63.25}{\sigma} = Z_{0.25} = 0674, \ \therefore \sigma = 13.442$

(3) $Q_1 = \mu - Z_{0.25}\,\sigma = 63.25 - 0.674 \times 13.442 = 54.19$

(2) $P(X > 80) = P\left(Z > \frac{80 - 63.25}{13.442}\right) = P(Z > 1.25) = 0.1056$

$\rightarrow 4000 \times 0.1056 = 422.4 = 422$ (人)

案例4

補6.10《台北大財政》

設有一銀行，其平時的營運狀況，平均3分鐘來一個顧客，請問：

(1)某日開始該銀行開始營業後，6分鐘內等不到任何一個客戶的機率是多少？

(2)第一位客人會在8～12分鐘內進入銀行的機率是多少？

(3)再請問：另一日，該銀行開始營業後，會在10分鐘內等到5位客人的機率是多少？

(4)又，請問：平時該銀行每日開始營業後，平均需等多久，才能等到前5個客人進入銀行？

案例4解說1

(1) 令X表示6分鐘內顧客到達數 → X ~ Poi(2) ，故

$$P(X = 0) = \frac{e^{-2}2^0}{0!} = e^{-2}$$

波松分配 Poisson distribution	在一個單位時段或區段內，某事件發生次數（X）的問題	$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0,1,2\ldots$ λ 為平均次數

(2)令T表示等待第1位顧客到達的時間，則

$$f(t) = \frac{1}{3}e^{-\frac{t}{3}} \quad t>0$$

指數分配 exponential distribution	與Poisson隨機變數相反，指數分配的隨機變數（X）是描述連續兩事件發生的間隔時間	$f(x) = \frac{1}{\beta}e^{-\frac{x}{\beta}}, x \geq 0$ x 為第一次發生事件所需時間； β 為事件發生的平均時間；

故第一位顧客會在8～12分鐘進入銀行之機率為

$$P(8 \leq T \leq 12) = \int_8^{12} \frac{1}{3}e^{-\frac{t}{3}}\,dt = \left[-e^{-\frac{1}{3}t}\right]_8^{12} = e^{-\frac{8}{3}} - e^{-4}$$

案例4解說2

(3)令X表示10分鐘內顧客到達人數，則 $X \sim Poi\left(\frac{10}{3}\right)$

故 $P(X=5)=\frac{e^{-\frac{10}{3}}\left(\frac{10}{3}\right)^5}{5!}$

(4)令Y表示等待第5位顧客到達的時間，則 $Y \sim \Gamma\left(\alpha=5, \lambda=\frac{1}{3}\right)$

故知平均等候時間為

$E(Y)=\frac{\alpha}{\lambda}=\frac{5}{1/3}=15$

註：λ = 1/β

Gamma 分配 Gamma distribution	隨機變數 (*X*) 表示事件第 *a* 次發生所需的時間。因此若「*X*>*t*」表示事件第 *a* 次發生至少需要 *t* 個時間單位；另個說法是，在 *t* 時間內事件至少發生 *(a-1)* 次的機率	$f(x)=\frac{1}{\beta\Gamma(a)}(\frac{x}{\beta})^{a-1}e^{-\frac{x}{\beta}}$, $0<x<\infty, a>0, \beta>0$ x 為第 *a* 次發生所需時間； β 表發生一次所需的時間；

Sampling Distributions

抽樣分配

案例5

補7.1【中山財管】
假設元富證券之股價X為一隨機變數，並有下列之隨機分配：

X	12	13	14
P(X=x)	0.4	0.4	0.2

(1)試求隨機抽樣本為2時，樣本平均數 $\bar{x}$ 之抽樣分配。

(2)根據(1)之抽樣分配試計算 E($\bar{x}$) 與 σ($\bar{x}$) 。

案例5解說

列出各種組合，記得驗證所算出項目的分配機率，
其累加總和應為「1」

(1)

$\overline{X}$	12	12.5	13	13.5	14	o.w.
$f_{\overline{X}}(\overline{x})$	0.16	0.32	0.32	0.16	0.04	0

(2) $E(\overline{X}) = 12.8$，$SD(\overline{X}) = 0.5292$

利用期望值與變異數的定義，可求得答案：

$$E(x) = \sum x\, f(x) \qquad Var(x) = \sigma^2(x) = E(X^2) - [E(X)]^2$$

案例6

補7.6《淡江企管》

淡江公司有4位員工，其薪資分別為2、3、3、4（萬元），

(1)現自其中以不放回的方式隨機抽出兩位員工的薪資，試分別求出員工平均薪資 $\bar{x}$ 及薪資變異數S^2 的抽樣分配。

(2)試分別求算 $\bar{x}$ 及 S^2 的期望值及變異數。

案例6解說1

(1)因採不放回方式，故知共有$\binom{N}{n}=\binom{4}{2}=6$組可能樣本，又

列出各種組合

(x_1,x_2)	$\bar{x}$	$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$
(2,3)	2.5	0.5
(2,3)	2.5	0.5
(2,4)	3	2
(3,3)	3	0
(3,4)	3.5	0.5
(3,4)	3.5	0.5

故知 $\bar{X}$ 之抽樣分配為

$\bar{x}$	2.5	3	3.5
$f(\bar{x})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

且 S^2 之抽樣分配為

s^2	0	0.5	2
$f(s^2)$	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$

統合所列出上述各種組合，
將之表列出來，即為
「抽樣分配」

案例6解說2

(2) $E(\bar{X}) = \sum \bar{x} f(\bar{x}) = 2.5 \times \frac{1}{3} + 3 \times \frac{1}{3} + 3.5 \times \frac{1}{3} = 3$

$Var(\bar{X}) = E(\bar{X}^2) - \left(E(\bar{X})\right)^2$

$= (2.5)^2 \times \frac{1}{3} + (3)^2 \times \frac{1}{3} + (3.5)^2 \times \frac{1}{3} - (3)^2 = \frac{1}{6}$

$E(S^2) = 0 \times \frac{1}{6} + 0.5 \times \frac{2}{3} + 2 \times \frac{1}{6} = \frac{2}{3}$

$Var(S^2) = E\left((S^2)^2\right) - \left(E(S^2)\right)^2$

$= (0)^2 \times \frac{1}{6} + (0.5)^2 \times \frac{2}{3} + (2)^2 \times \frac{1}{6} - \left(\frac{2}{3}\right)^2 = \frac{7}{18}$

$\mathrm{E(x)} = \sum \mathrm{x\,f(x)}$ $\mathrm{Var(x)} = \sigma^2(\mathrm{x}) = \mathrm{E}(X^2) - [\mathrm{E(X)}]^2$

Estimation and Interval Estimation

估計與區間估計

案例7

補8.5【台大國發】

假設X的母體分配為常態，其期望值為μ，變異數為σ^2，考慮從隨機樣本觀察值X_1，X_2，.....，X_n估計μ，以下為3個估計式：

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \text{，} \quad \hat{\mu}=\frac{1}{n+1}\sum_{i=1}^{n}x_i \text{，}$$

$$\bar{\mu}=\frac{1}{2}x_1+\frac{1}{2n}\sum_{i=2}^{n}x_i$$

試問此三個估計值是否滿足：

(1) 不偏性（unbiasedness）？

(2) 漸進不偏性（asymptotic unbiasedness）？

(3) 一致性（consistency）？

案例7解說1

(1) $E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n}n\mu = \mu$ （不偏）

$E(\hat{\mu}) = E\left(\frac{1}{n+1}\sum_{i=1}^{n} X_i\right) = \frac{1}{n+1}\sum_{i=1}^{n} E(X_i) = \frac{1}{n+1}\sum_{i=1}^{n} \mu = \frac{n}{n+1}\mu \neq \mu$ （偏誤）

$E(\tilde{\mu}) = E\left(\frac{X_1}{2} + \frac{\sum_{i=2}^{n} X_i}{2n}\right) = \frac{\mu}{2} + \frac{\sum_{i=2}^{n} \mu}{2n} = \frac{n\mu}{2n} + \frac{(n-1)\mu}{2n} = \frac{2n-1}{2n}\mu \neq \mu$ （偏誤）

利用滿足「不偏性」的定義，作為判斷的方式

$E(\hat{\theta}_n) = \theta$,

案例7解說2

(2) $\lim\limits_{n\to\infty} E(\overline{X}) = \lim\limits_{n\to\infty} \mu = \mu$ （漸近不偏）

$$\lim_{n\to\infty} E(\hat{\mu}) = \lim_{n\to\infty} \frac{n}{n+1}\mu = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}}\mu = \mu \text{（漸近不偏）}$$

$$\lim_{n\to\infty} E(\tilde{\mu}) = \lim_{n\to\infty} \frac{2n-1}{2n}\mu = \lim_{n\to\infty} \frac{2-\frac{1}{n}}{2}\mu = \mu \text{（漸近不偏）}$$

利用滿足「漸進不偏性」的定義，作為判斷的方式

$$\lim_{n\to\infty} E(\hat{\theta}_n) = \theta \text{,}$$

案例7解說3

(3) $\overline{X}, \hat{\mu}$ 與 $\tilde{\mu}$ 三者均已漸近不偏，只需再考慮各自變異數 $n \to \infty$ 時是否收斂為 0

$$\mathrm{Var}(\overline{X}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow \lim_{n\to\infty}\mathrm{Var}(\overline{X}) = \lim_{n\to\infty}\frac{\sigma^2}{n} = 0 \text{（具有一致性）}$$

滿足「一致性」的原始定義：

$$P\left(\lim_{n\to\infty}\hat{\theta}_n = \theta\right) = 1$$

$$\mathrm{Var}(\hat{\mu}) = \mathrm{Var}\left(\frac{1}{n+1}\sum_{i=1}^{n} X_i\right) = \frac{1}{(n+1)^2}\sum_{i=1}^{n}\sigma^2 = \frac{n}{(n+1)^2}\sigma^2$$

$$\Rightarrow \lim_{n\to\infty}\mathrm{Var}(\hat{\mu}) = \lim_{n\to\infty}\frac{n}{(n+1)^2}\sigma^2 = 0 \text{（具有一致性）}$$

滿足「一致性」的判斷：

$$\lim_{n\to\infty}\mathrm{Var}(\hat{\theta}_n) = 0$$

$$\mathrm{Var}(\tilde{\mu}) = \mathrm{Var}\left(\frac{X_i}{2} + \frac{\sum_{i=2}^{n} X_i}{2n}\right) = \frac{\sigma^2}{4} + \frac{\sum_{i=2}^{n}\sigma^2}{4n^2} = \frac{n^2\sigma^2}{4n^2} + \frac{(n-1)\sigma^2}{4n^2} = \frac{n^2+n-1}{4n^2}\sigma^2$$

$$\Rightarrow \lim_{n\to\infty}\mathrm{Var}(\tilde{\mu}) = \lim_{n\to\infty}\frac{n^2+n-1}{4n^2}\sigma^2 = \lim_{n\to\infty}\left(\frac{1}{4} + \frac{1}{4n} - \frac{1}{4n^2}\right)\sigma^2 = \frac{1}{4}\sigma^2 \neq 0 \text{（不具一致性}$$

案例8

補8.17《東吳國貿》

假設 μ_A，σ_A^2 及 μ_B，σ_B^2 依序為所有A型燈泡及B型燈泡的常態母體平均數及變異數，而且已知 $\sigma_A = 26$ 及 $\sigma_B = 27$。今隨機抽取A型燈泡40個及B型燈泡50個進行平均壽命檢驗，A型燈泡的平均壽命為418小時，B型燈泡的平均壽命為402小時。試求 $\mu_A - \mu_B$ 的95 %信賴區間。

案例8解說

$\mu_A-\mu_B$ 的95%信賴區間為

$$\left(\bar{X}_A-\bar{X}_B-z_{0.025}\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}},\bar{X}_A-\bar{X}_B+z_{0.025}\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}\right)$$

$$\Rightarrow\left(418-402-1.96\sqrt{\frac{(26)^2}{40}+\frac{(22)^2}{50}},418-402+1.96\sqrt{\frac{(26)^2}{40}+\frac{(22)^2}{50}}\right)$$

$$\Rightarrow(5.895,26.105)$$

$$\bar{x}_1-\bar{x}_2\sim N\left(\mu_1-\mu_2\ ,\ \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$$

案例9

補8.19《台大商研》

(1) 以90%的信賴水準且誤差範圍在0.03之內，請問估計一個母體比例值所需要的樣本大小是多少？假設樣本比例值未知。

(2) 假設你知道樣本比例值應該不會小於0.75，請重做(1)小題．

(3) 假設使用(2)的計算結果得到樣本大小．發現樣本比例值為0.92．請以90 %的信賴水準估計母體比例值。

案例9解說

樣本比例未知，所以用 p=1/2估算

(1) $n=\frac{1}{4}\left(\frac{z_{0.05}}{d}\right)^2=\frac{1}{4}\left(\frac{1.645}{0.03}\right)^2=751.67$，取 $n=752$

利用所給的樣本比例值來估算

(2) $n=\left(\frac{z_{0.05}}{d}\right)^2(0.75)(0.25)=563.755$，取 $n=564$

(3) 因 $\hat{p}=0.92$，故 p 之90%信賴區間為

$$\Rightarrow\left(\hat{p}\mp z_{0.05}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)=\left(0.92\mp 1.645\cdot\sqrt{\frac{0.92\times 0.08}{564}}\right)$$

$$=(0.9012,0.9388)$$

觀念

《淡江管科、清大工工》

Suppose that we obtain a 95% confidence of the mean μ to be (65.5, 68.4). We know that $P(65.5 \leq \mu \leq 68.4) = 0.95$.

「錯」，看這句〝P（...）=.......〞的意義，μ的機率不是0，就是1

The End