

Probability and Statistics, Spring 2018

Problem Design 2

DUE DATE: May 28, 2018

第 4 組

Answer

1. X 是一個高斯是一個高斯隨機變數，且 $E[X] = 0$, $P[|X| \leq 10] = 0.1$ ，試問 X 的標準差為多少？

$$P[|X| \leq 10] = P[-10 < X < 10] = \Phi\left(\frac{10}{\sigma}\right) - \Phi\left(\frac{-10}{\sigma}\right) = 2\Phi\left(\frac{10}{\sigma}\right) - 1 = 0.1.$$

$$\Phi\left(\frac{10}{\sigma}\right) = 0.55 \rightarrow \frac{10}{\sigma} \approx 0.13, \sigma \approx 76.92.$$

2. 羅斯福路和基隆路口，南北向的號誌燈紅燈每 110 秒、綠燈 90 秒交錯一次。任一位行人在紅燈時停下。

1. (1) 至少等待 1 分鐘以上才能過馬路的機率為何？

$$P[X \geq 1\text{min}] = \frac{5}{11}.$$

- (2) 等待時間介於半分鐘和 1 分半鐘的機率為何？

$$P[0.5\text{min} \leq X \leq 1\text{min}] = \frac{3}{11}.$$

2. X 是一個高斯隨機變數，而今天有一個人每天喝的手搖杯數也是高斯分配 ($\mu = 5, \sigma = 1$)，求 $P[4 < X \leq 6] = ?$

$$P[4 < X \leq 6] = \Phi\left(\frac{6-5}{1}\right) - \Phi\left(\frac{4-5}{1}\right) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0.6826.$$

3. A 公司每被客户提起一次訴訟，理論上就必須提供客戶賠償。但因為 A 公司的律師團隊運作良好，因此在 1000 次的訴訟中，只有 25 次需要賠償超過或等於 668 元美金的賠款。而 A 公司付出的賠償金額近似於 σ 為 200 的高斯隨機變數。

客戶在提起訴訟時，同樣也會需要支付訴訟的費用。普遍來說，法官認為客戶需要付出 200 美金的訴訟費用。

請問，當提起一次訴訟時，客戶有多少機率是不會因此而虧本的？

$$P[X \geq 668] = 1 - P[X < 668] = \frac{25}{1000} = 1 - \Phi\left(\frac{668-\mu}{200}\right).$$

$$\Phi\left(\frac{668-\mu}{200}\right) = \frac{975}{1000} \rightarrow \frac{668-\mu}{200} = 1.96, \mu = 276.$$

$$P[X \geq 200] = 1 - P[X < 200] = 1 - \Phi\left(\frac{200-276}{200}\right) = \Phi(0.38) = 0.648.$$

5. 台中五月份的溫度，以攝氏計，為高斯隨機變數，變異數為 10，溫度超過 27 度的機率為二分之一。請問當溫度大於 40 度以及溫度小於 24 度時，的機率分別為多少？

$$P[X \geq 27] = \frac{1}{2} = 1 - \Phi\left(\frac{27-\mu}{\sqrt{10}}\right), \Phi\left(\frac{27-\mu}{\sqrt{10}}\right) = \frac{1}{2} \rightarrow \mu = 27.$$

$$P[X \geq 40] = 1 - \Phi\left(\frac{40-27}{\sqrt{10}}\right) \approx Q(4.11) = 1.98 \cdot 10^{-5}.$$

$$P[X < 24] = \Phi\left(\frac{24-27}{\sqrt{10}}\right) = 1 - \Phi\left(\frac{2}{\sqrt{10}}\right) \approx 1 - 0.8289 = 0.1711.$$

6. 某電子商品零售店其某品牌數位相機之每週銷售量為常態分佈其平均值為 100 與標準偏差為 10，此零售店每週一進貨，為了要保證每週只有 5% 缺貨的機會，零售店經理週一要進貨幾架此品牌的數位相機？

假設 K 為至少幾架數位相機，

$$P[X > K] = 0.05 \rightarrow 1 - \Phi\left(\frac{K-100}{10}\right) = 0.05 \rightarrow \Phi\left(\frac{K-100}{10}\right) = 0.95, \frac{K-100}{10} = 1.645.$$

$$K = 116.45 \rightarrow 117 \text{ 架 (整數)}$$

7. X 是一個 PDF 為 $f(t)$ 的正連續隨機變數，如果 $g(t) = tf(t)$ 同樣為一隨機變數 Y 的 PDF, $E(X)$ 的值為何？並且使用 $E(Y)$ 來表達 $\text{Var}(x)$ 。

$$E[X] = \int_{-\infty}^{\infty} tf(t)dt.$$

$$E[Y] = \int_{-\infty}^{\infty} t^2 f(t)dt.$$

$$\text{Var}[X] = E[X^2] - (E[X])^2 = E[Y] - (E[X])^2$$

8. X 和 Z 是獨立的隨機變量, $E[X] = E[Z] = 0$, $\text{Var}[X] = 1$, $\text{Var}[Z] = 16$ 。現在令 $Y = X + Z$ 。找出 X 和 Y 的相關係數 $\rho_{X,Y}$, X 和 Y 是獨立的嗎？

$$E[Y] = E[X + Z] = E[X] + E[Z] = 0.$$

$$\text{Var}[Y] = \text{Var}[X + Z] = \text{Var}[X] + \text{Var}[Z] = 1 + 16 = 17, \text{Cov}[X, Z] = 0 \quad (X, Z \text{ 獨立})$$

$$\text{Cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[(X - \mu_X)(X + Z - \mu_X - \mu_Z)] = \text{Var}[X] + \text{Cov}[X, Z] = 1 + 0 = 1.$$

$$\rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y} = \frac{1}{\sqrt{1 \cdot 17}} \approx 0.24.$$

Ans: X, Y 不獨立, 因為 $\text{Cov}[X, Y] \neq 0$.

9. 甲說實話的機率為 $7/10$, 乙說實話的機率為 $9/10$, 今有一袋內藏 3 白球, 7 黑球, 自袋中任取一球, 甲乙二人均說是白球, 則此球確實為白球之機率為何？

$$P[\text{白球}] = \frac{9 \cdot 7 \cdot 3}{9 \cdot 7 \cdot 3 + 3 \cdot 1 \cdot 7} = \frac{189}{210} = \frac{9}{10}.$$

10. 今天鬍子哥的健身房的槓片標示 45 磅, 槓子實為 45 磅, 槓片實際重量為高斯分佈, $\mu = 44$ 磅, $\sigma = 2$ 。已知兩邊重量相差 10%, 使用者將會受傷。鬍子哥今天將深蹲 135 磅 (一邊各一片), 請問有多少機率會受傷？(Hint: 已經確認一邊槓片實重 45 磅。槓片邊界有無包含沒有關係, 以好算為準)

$$P[\text{受傷}] = P[\text{另一槓片} \geq 49.5] + P[\text{另一槓片} \leq 40.5]$$

$$P_X[X \geq 49.5] = 1 - \Phi\left(\frac{49.5 - 44}{2}\right) = 1 - \Phi(2.75) = 1 - 0.99702 = 2.98 \cdot 10^{-3}.$$

$$P_X[X \leq 40.5] = \Phi\left(\frac{40.5 - 44}{2}\right) = \Phi(-1.75) = 1 - \Phi(1.75) = 1 - 0.9599 = 0.0401$$

$$P[\text{受傷}] = 2.98 \cdot 10^{-3} + 0.0401 = 0.04308.$$

11. 設筒中有 1 號球 1 個, 2 號球 2 個, , N 號球 N 個, 每球被抽中的機率相等, 若抽到 K 號球能得 K 元, 試求:

1. 抽一顆球抽到 K 號球的機率為？其中 K 為整數且 $1 \leq K \leq N$.

$$P[K] = \frac{K}{1+2+\dots+N} = \frac{2K}{N(1+N)}.$$

2. 抽一顆球的期望值和變異數為？

$$E[K] = K \cdot P[K] = \frac{2K^2}{N(1+N)}.$$

$$\text{Var}(K) = E[K^2] - (E[K])^2 = \frac{2K^3}{N(1+N)} - \left(\frac{2K^2}{N(1+N)}\right)^2 = \frac{2K^2}{N(1+N)} \left[K - \frac{2K^2}{N(1+N)}\right].$$

12. 假設今天你參與一個遊戲, 關主的桌上有三個倒立蓋著的杯子, 只有一個杯子蓋著獎品, 而關主已經知道哪個杯子蓋著獎品, 關主會先要你猜哪個杯子裡有獎品, 當你選定其中一個之後, 關主會故意在另外兩個杯子中, 翻開一個他知道沒有獎品的杯子, 接下來他會問你要不要更換你原先的選擇, 選剩下另一個沒被翻開的杯子, 請問你是要換還是不要換？原因是為什麼？

有三種可能的情況, 全部都有相同機率 ($1/3$):

- 你選到獎品, 關主挑兩空杯子任一杯, 換的話將失敗。
- 你選到空杯 A, 關主挑空杯 B, 換的話將得到獎品。
- 你選到空杯 B, 關主挑空杯 A, 換的話將得到獎品。

13. 投擲一個公正的硬幣。如果是硬幣結果是正面 (head)，則投擲一個公正的六面骰子 100 次；如果是硬幣結果是反面 (tail)，則投擲一個公正的六面骰子 101 次。令 X 事件代表投擲硬幣的結果，正面為 1(head) 反面為 0(tail)；令 Y 事件代表骰子擲出 6 點的次數，請求出 $P(X = 1 | Y = 15)$

$$\frac{\frac{\binom{100}{15}}{6^{100}}}{\frac{\binom{100}{15}}{6^{100}} + \frac{\binom{101}{15}}{6^{101}}}.$$

14. $P_{N,K}(n, k) = \begin{cases} \frac{(1-p)^{n-1}p}{n}, & k = 1, 2, \dots, n; n = 1, 2, \dots, \\ 0, & \text{otherwise.} \end{cases}$

請找出 maginal PMF $P_N(n)$, $E[N]$, $\text{Var}[N]$, $E[N^2]$, $E[K]$, $\text{Var}[K]$, $E[N + K]$, $E[NK]$, $\text{Cov}[N, K]$ 。

The marginal PMF of N . For $n = 1, 2, \dots$,

$$P_N(n) = \sum_{k=1}^n \frac{(1-p)^{n-1}p}{n} = (1-p)^{n-1}p.$$

We can obtain that N has a geometric PMF. Thus,

$$E[N] = \frac{1}{p} \quad \text{Var}[N] = \frac{1-p}{p^2}.$$

$$E[N^2] = \text{Var}[N] + (E[N])^2 = \frac{2-p}{p^2}.$$

$$E[K] = \sum_{n=1}^{\infty} \sum_{k=1}^n k \frac{(1-p)^{n-1}p}{n} = \sum_{n=1}^{\infty} \frac{(1-p)^{n-1}p}{n} \sum_{k=1}^n k = \frac{1}{2p} + \frac{1}{2}.$$

$$E[K^2] = \sum_{n=1}^{\infty} \sum_{k=1}^n k^2 \frac{(1-p)^{n-1}p}{n} = \sum_{n=1}^{\infty} \frac{(1-p)^{n-1}p}{n} \sum_{k=1}^n k^2 = \frac{2}{3p^2} + \frac{1}{6p} + \frac{1}{6}.$$

$$\text{Var}[K] = E[K^2] - (E[K])^2 = \frac{5}{12p^2} - \frac{1}{3p} + \frac{5}{12}.$$

$$E[N + K] = E[N] + E[K] = \frac{3}{2p} + \frac{1}{2}.$$

$$E[NK] = \sum_{n=1}^{\infty} \sum_{k=1}^n nk \frac{(1-p)^{n-1}p}{n} = \sum_{n=1}^{\infty} (1-p)^{n-1}p \sum_{k=1}^n k = \frac{1}{p^2}.$$

$$\text{Cov}[N, K] = E[NK] - E[N]E[K] = \frac{1}{2p^2} - \frac{1}{2p}.$$

15. 精華商場買進一批超薄筆電，平均進價每台 18000 元，標準差 1000 元，今商場擬依進價的 150% 再加 2000 元為售價，試問這批超薄筆電的平均售價每個多少元？又其標準差為多少？

$$E[\text{price}] = 18000 \cdot 150\% + 2000 = 29000.$$

$$\sigma_{\text{price}} = 1000 \cdot 150\% = 1500.$$

18. 人豪和國豪是在機率課認識的同學，有天他們都因為同時沒趕上公車而在公車站牌處巧遇，人豪很擔心上課就要遲到了，所以非常焦慮，這時候國豪拍拍她的肩膀並安慰他說：「沒關係啦！你知道嗎？其實等公車的過程就是一種 Poisson 分佈阿！」人豪紅著眼眶說：「我知道阿，所以呢？」國豪眼看表現的機會來了，立刻回想昨天跟同學借來抄的機率作業答案，然後接著說：「如果我們現在開始等 10 分鐘，等到一台公車的機率有 0.303 呢！」國豪聽完當場臉綠掉，接著說：「哼！看來作業是你自己寫的對吧」

- (1) 如果國豪說的沒錯，那請問平均多久會等到一班公車？人豪當場心算後說：「那照你這個邏輯，10分鐘之內等不到公車的機率反而有 0.607 呢！」國豪聽完當場臉綠掉，接著說：「哼！看來作業是你自己寫的對吧！那我就來考考你。」

$$P[\text{等到一台公車}] = 0.303 = \frac{\alpha^x e^{-\alpha}}{x!} = \frac{\alpha^1 e^{-\alpha}}{1!} \rightarrow \alpha = 1.758.$$

$$\lambda = \frac{1.758}{10} = 0.1758 \rightarrow \frac{1}{0.1758} = 5.6882 \text{ (min / hit)}.$$

- (2) 這號公車平均每 12 分鐘會來一班，請問在 20 分鐘內，等不到公車和等到 1 班公車的機率分別為多少？

$$\lambda = \frac{1}{12}, T = 20, \alpha = \frac{5}{3}.$$

$$P[\text{等不到公車}] = e^{-5/3}.$$

$$P[\text{等到 1 班公車}] = \frac{5}{3} e^{-5/3}.$$

- (3) 承上題，請問在半小時內，等到車的機率為多少？

$$\lambda = \frac{1}{12}, T = 30, \alpha = \frac{5}{2}$$

$$P[\text{在半小時內等到車}] = 1 - e^{-5/2}.$$

- (4) 承上題，假設等計程車的過程中為指數分佈，而在 18 分鐘內（包含）等到一輛計程車的機率為 0.835，另外，計程車與公車到達目的地的乘車時間相同，請問人豪與國豪應該選擇搭乘哪一種交通工具為優？

$$f_X(x) = \lambda e^{-\lambda x} x \geq 0$$

$$\lambda = \frac{1}{18}$$

$$\lambda \int_0^{18} e^{-\lambda x} dx = 0.835 \rightarrow \lambda = 0.1.$$

平均 10 分鐘來一輛計程車較可能。

- (5) 請問 Poisson 分佈與指數分佈相同嗎？如果不同，那他們之間有什麼關係嗎？

不同，Poisson 是針對事件發生次數的機率，是離散的，指數是針對時間，是連續的，兩者有類似「倒數」之關係。

- (6) 請問等公車的過程可以視為一個均等分佈嗎？請解釋為什麼或者需要什麼前提假設嗎？

不太行，從 exponential 的角度看，等到 $t \rightarrow \infty$ 則車一定會來，機率 $\rightarrow 1$ ，一定等的到，均勻分佈是在一個時間區間 (a, b) 若 $b \rightarrow \infty$ 則 $\frac{1}{b-a} \rightarrow 0$ ，無意義，因此等公車模型不適合用 Uniform Random Variable 描述。

在討論與解完這些問題後，志明與志朋決定搭乘他們所討論出來的交通工具，但是就在他們說完話的同時，迎面而來的是 出現機率較低的交通工具，於是 他們就上車了。

Ranking

- (1) 本題要問的是 Gaussian Random Variables 相關計算
- (2) 透過絕對值的限制，可以導出一個對稱的作圖
- (3) 題目設計蠻直接，沒有一些應用的想法，可以為此題目設想一些情境。
- (1) 本題要問的是 Uniform 與 Gaussian Random Variables 的應用計算
- (2) 題目很生活化，透過我們熟悉的過馬路與喝手搖杯生活經驗練習高斯分配。
- (3) Uniform Random Variables 的部分有點太簡單，憑直覺就可以寫出答案，可以設計成需要一些計算才能得到答案。
- (1) 本題要問的是 Gaussian Random Variables 的應用計算。
- (2) 題目有經過設計，根據少許的訊息和限制，可以找到整個高斯函數的模型。
- (3) 最後算出來的數字好像有違於題目所說的律師團隊運作良好，可以將客戶虧本數字設計的較小。

5. (1) 除了 Gaussian Random Variables 的 Standard normal distribution, 也要用到 standard normal complementary CDF。
(2) 先前機率設為 $1/2$, 可以考的觀念在於高斯分布中間值極為 $1/2$ 。
(3) 蠻直接即能回答, 或許可以將數字設計在一個標準差上, 能夠考更多觀念。
6. (1) 本題要問的是 Gaussian Random Variables 的應用計算
(2) 題目設計很好, 商業應用到電商零售業的存貨管理上, 用高斯分布算進貨量以控制缺貨機率, 很實用。
(3) 出的很好令人驚艷, 沒什麼可以挑剔的。
7. (1) 本題要問的是函數變數的期望值與變異數。
(2) 讓人能從不同角度思考這類型題目。
(3) 題意有點不清楚, 可以再明確一點。
8. (1) 本題要問的是 Correlation Coefficient 的相關計算。
(2) 經典題型, 可讓人熟悉這單元, 出的不錯。
(3) 經典, 但有點太簡單, 依照這種題型 SOP 就可以直接解出, 如果可以想出一些變化或不同前題會比較有趣。
9. (1) 本題要問的是 Bayes' Theorem 的相關計算。
(2) 經典貝氏定理題型, 可讓人熟悉這單元, 出的不錯。
(3) 經典題型, 如果要增加難度或變化的話, 可以增加人數, 或增加球數。
10. (1) 高斯分布的應用、集合的聯集。
(2) 與生活結合題目內容有趣, 高斯分布不錯的應用。
(3) 其實根本跟槓子本身沒太大的關係, 如果要設計槓子, 可以加入槓子的重量也是高斯分布, 探討兩個高斯分布關係增加題目豐富度。
11. (1) 排列組合的應用與期望值、變異數之定義。
(2) 題幹敘述清楚言簡意賅。
(3) 有點太單調, 可能可以多抽幾個球。
12. (1) 條件機率。
(2) 三門問題變成三個杯子問題
(3) 幾乎完全沒有改三門問題...上次團體作業已經有許多類似並加入新元素的題目了, 不知道這組這次為啥竟然出了一個完全沒有創新的題目。可以改成多們多禮物之類的都比這原來網路上照抄的好。
13. (1) 條件機率、多變數隨機變數。
(2) 結合了多變數、條件機率的概念。
(3) 有利用新學概念與舊有的排列組合條件機率、算是不錯的題目了, 不過題目的敘述應該改成 $P[X = 1 | Y = 15]$, 中括號 (這其實是很小很小的問題啦)。
14. (1) 多變數隨機變數的期望值、變異數、Cov、marginal PMF 的關係
(2) 題目敘述有趣 (?) 硬是編了一個小故事 XD。
(3) 結合應用, 把小故事改成真的是這個分布的情況, 在作答的過程可以發現 marginal PMF 根本是 Geometrical distribution, 應該有許多實際的例子可以套, 而不是純粹硬掰一個小故事 (雖然也滿好笑 XD)

15. (1) Derived Random Variable。
(2) 生活化的數學問題。
(3) 有點過度簡單，可以加入一些多變數隨機變數的概念。
18. (1) 波松分布、指數分布。
(2) 十分生活化，題目十分有趣而且在數學的部分也有深層的了解，確實有讓人重新思考波松分布與指數分布的關係，同時針對這兩個分布十分良好的生活應用。
(3) 幾乎沒什麼好挑剔的了，完全就是應用這堂機率課的概念解決生活的問題。