



# 机率

台大电机系 叶丙成

微博: [weibo.com/yehbo](https://weibo.com/yehbo) 脸书: [facebook.com/prof.yeh](https://facebook.com/prof.yeh)

部落格: [pcyeh.blog.ntu.edu.tw](http://pcyeh.blog.ntu.edu.tw)



Prof. Yeh, Ping-Cheng (Benson) 葉丙成  
Dept. of EE, National Taiwan University

# 本周主题概述

---

- 7-1: 期望值 II
- 7-2: 随机变量之函数
- 7-3: 条件机率分布与失忆性





# 7-1: 期望值 II (EXPECTATION)

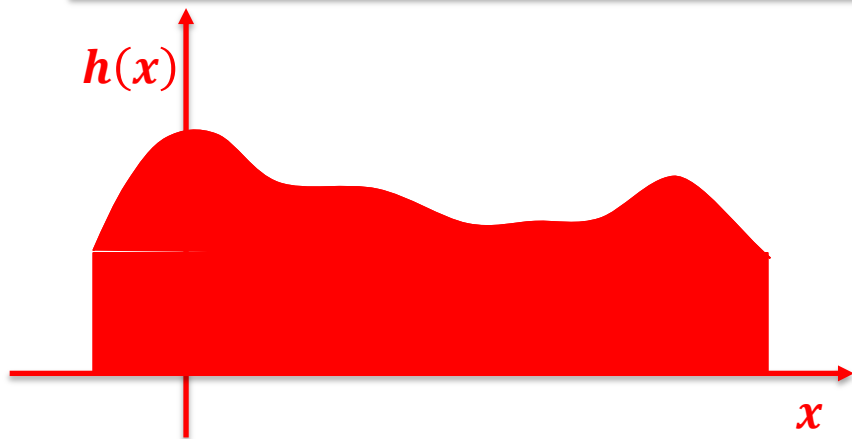
---

第七周

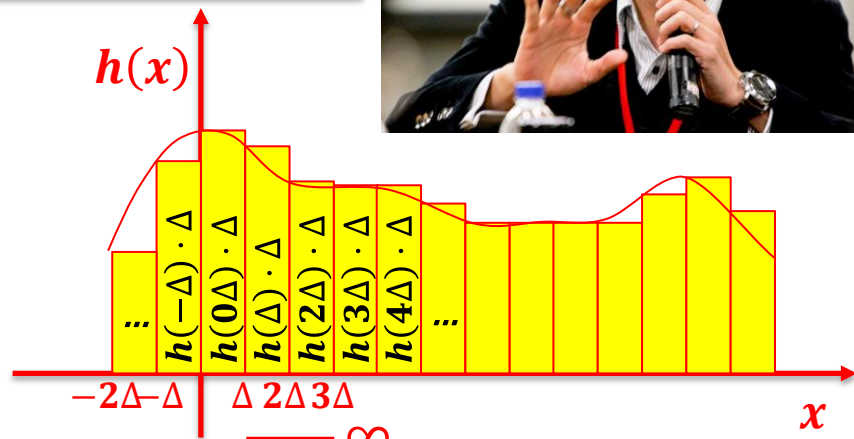


Prof. Yeh, Ping-Cheng (Benson) 葉丙成  
Dept. of EE, National Taiwan University

# 积分的近似概念



$$\text{面积} = \int_{-\infty}^{\infty} h(x) dx$$



$$\text{面积} = \sum_{n=-\infty}^{\infty} h(n\Delta) \cdot \Delta$$

$$\Rightarrow \text{面积} = \int_{-\infty}^{\infty} h(x) dx = \lim_{\Delta \rightarrow 0} \sum_{n=-\infty}^{\infty} h(n\Delta) \cdot \Delta$$



# 期望值 (Expectation)

- 对连续的随机变数  $X$  而言，想求期望值，我们用类似离散随机变量的方式出发
- 将  $X$  的值以  $\Delta$  为单位无条件舍去来近似  
结果：离散随机变数  $Y$  （当  $\Delta \rightarrow 0$  时， $X \approx Y$ ）
- 根据第五周：

$$- p_Y(n\Delta) = P(n\Delta \leq X < n\Delta + \Delta) \approx f_X(n\Delta) \cdot \Delta$$

$$\begin{aligned} X \in [0, \Delta) &\rightarrow Y = 0\Delta \\ X \in [1\Delta, 2\Delta) &\rightarrow Y = 1\Delta \\ &\vdots \end{aligned}$$

$$X \in [n\Delta, (n+1)\Delta) \rightarrow Y = n\Delta$$

$$\begin{aligned} E[X] &= \lim_{\Delta \rightarrow 0} E[Y] = \lim_{\Delta \rightarrow 0} \sum_{n=-\infty}^{\infty} n\Delta \cdot P_Y(n\Delta) \\ &= \lim_{\Delta \rightarrow 0} \sum_{n=-\infty}^{\infty} n\Delta \cdot f_X(n\Delta) \cdot \Delta = \int_{-\infty}^{\infty} x f_X(x) dx \end{aligned}$$



# 随机变量的函数之期望值



- 对于任一连续随机变数  $X$  而言，其任意函数  $g(X)$  亦是一随机变量，亦有期望值
- $g(X)$  期望值定义为

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx$$

※ 离散随机变数： $E[g(X)] = \sum_{x=-\infty}^{\infty} g(x)p_X(x)$



# 期望值运算的性质



- $E[\alpha g(X) + \beta h(X)]$

$$= \int_{-\infty}^{\infty} [\alpha g(x) + \beta h(x)] \cdot f_X(x) dx$$

- $= \alpha \cdot \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx + \beta \cdot \int_{-\infty}^{\infty} h(x) \cdot f_X(x) dx$

$$= \alpha \cdot E[g(X)] + \beta \cdot E[h(X)]$$

Ex:  $E[6X + 8X^2] = 6E[X] + 8E[X^2]$



# 期望值运算的性质

- $E[\alpha]$

$$= \int_{-\infty}^{\infty} \alpha \cdot f_X(x) dx = \alpha \cdot \int_{-\infty}^{\infty} f_X(x) dx = \alpha$$

Ex:  $E[6] = 6$





# 常见的随机变量函数期望值



- $X$  的  $n^{th}$  moment :

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

– Ex:  $E[X^2]$  是  $X$  的  $2^{nd}$  *moment*

– Ex:  $E[X^5]$  是  $X$  的  $5^{th}$  *moment*

- $X$  的 变异数 (variance) :

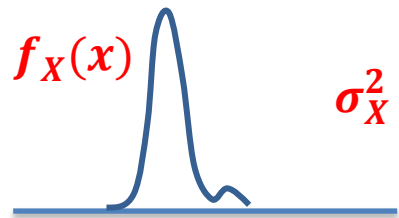
$$E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_x)^2 f_X(x) dx$$



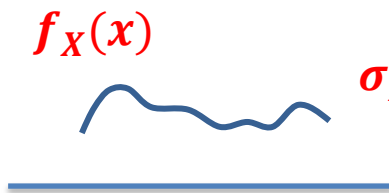
# 变异数 (Variance)



- Variance 通常符号表示为  $\sigma_X^2 = E[(X - \mu_X)^2]$
- 变异数隐含关于随机变数  $X$  多「乱」的信息



$\sigma_X^2$  小 ( $\because x \approx \mu_X$ )



$\sigma_X^2$  大 ( $\because X$  不见得接近  $\mu_X$ )

- 变异数的开根号便是标准差 (standard deviation) :  $\sigma_X$

$$\parallel \\ \sqrt{\text{Variance}}$$



# Variance 便利算法



$$\begin{aligned}\bullet \sigma_X^2 &= E[(X - \mu_X)^2] \\ &= E[X^2 - 2\mu_X \cdot X + \mu_X^2] \\ &= E[X^2] + E[-2\mu_X X] + E[\mu_X^2] \\ &= E[X^2] - \underbrace{2\mu_X \cdot E[X] + \mu_X^2}_{-2\mu_X^2 + \mu_X^2} = \boxed{E[X^2] - \mu_X^2} \\ &\Rightarrow E[X^2] = \sigma_X^2 + \mu_X^2\end{aligned}$$



# 常见连续分布之期望值/变异数

- $X \sim \text{Exponential}(\lambda)$  :

- $\mu_X = \frac{1}{\lambda}$

- $\sigma_X^2 = \frac{1}{\lambda^2}$

- $X \sim \text{Erlang}(n, \lambda)$  :

- $\mu_X = \frac{n}{\lambda}$

- $\sigma_X^2 = \frac{n}{\lambda^2}$



# 常见连续分布之期望值/变异数

- $X \sim \text{Gaussian}(\mu, \sigma)$  :

- $\mu_X = \mu$

- $\sigma_X^2 = \sigma^2$

- $X \sim \text{UNIF}(a, b)$  :

- $\mu_X = \frac{a+b}{2}$

- $\sigma_X^2 = \frac{1}{12}(b-a)^2$



# 期望值推导范例

分部积分：

$$\int U dV = UV - \int V dU$$

- $X \sim \text{Exponential}(\lambda) : f_X(x) = \lambda e^{-\lambda x}, x \geq 0$

$$\begin{aligned}\mu_X &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx \\&= - \int_0^{\infty} -x \lambda e^{-\lambda x} dx = - \int_0^{\infty} x e^{-\lambda x} d(-\lambda x) = - \int_0^{\infty} \underbrace{x}_U d \underbrace{e^{-\lambda x}}_V \\&= - \left[ x e^{-\lambda x} \Big|_0^{\infty} - \int_0^{\infty} e^{-\lambda x} dx \right] \\&= - \left[ 0 - 0 - \int_0^{\infty} e^{-\lambda x} dx \right] = \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda} \int_0^{\infty} \lambda e^{-\lambda x} dx \\&= \frac{1}{\lambda}\end{aligned}$$



# 本节回顾

---

- 连续随机变数的期望值定义？
- 连续随机变量的函数的期望值？
- 「凑」字诀！
- 常见连续机率分布之期望值、变异数？





# 7-2: 随机变量之函数

---

第七周



Prof. Yeh, Ping-Cheng (Benson) 葉丙成  
Dept. of EE, National Taiwan University



# 随机变数的函数

---

- 随机变数  $X$  的任意函数  $g(X)$  也是一个随机变数
- 通常被称为 Derived Random Variable



# 如何求 $g(X)$ 机率分布？

---

- 若  $X$  为离散：
  - 直接推  $g(X)$  的 PMF
- 若  $X$  为连续：
  - 先推  $g(X)$  的 CDF，再微分得到 PDF



# 离散 $X$ 之函数



- Ex: 某宅宅超爱战LOL。每次一战就连续战  $X$  场不可收拾，已知  $X \sim \text{GEO}(0.2)$ 。某宅宅内心仍有一点清明，其良心亦会因战过度而内疚，依战的次数多寡，内疚程度  $Y$  分别为 1, 2, 3 不同等级:

$$Y = g(X) = \begin{cases} 1, & \text{if } 1 \leq X \leq 3; \\ 2, & \text{if } 4 \leq X \leq 6; \\ 3, & \text{if } X \geq 7. \end{cases}$$

问  $Y = g(X)$  的机率分布？



# 离散 $X$ 之函数



- $Y = g(X) = \begin{cases} 1, & \text{if } 1 \leq X \leq 3; \\ 2, & \text{if } 4 \leq X \leq 6; \\ 3, & \text{if } X \geq 7. \end{cases}$
- $X \sim GEO(0.2) \Rightarrow p_X(x) = (1 - 0.2)^{x-1} \cdot 0.2$
- $p_Y(1) = p_X(1) + p_X(2) + p_X(3)$   
 $= 0.2 + 0.8 \cdot 0.2 + 0.8^2 \cdot 0.2$
- $p_Y(2) = p_X(4) + p_X(5) + p_X(6)$   
 $= (0.8)^3 \cdot 0.2 + (0.8)^4 \cdot 0.2 + (0.8)^5 \cdot 0.2$
- $p_Y(3) = P(Y = 3) = 1 - p_Y(1) - p_Y(2)$



# 离散 $X$ 之函数

- 当  $X$  为离散随机变数时，  
 $Y = g(X)$  亦为离散随机变数
- $Y = g(X)$  之 PMF 为

$$p_{g(X)}(y) = \sum_{\substack{\text{会让} \\ g(x)=y \\ \text{的所有 } x}} p_X(x)$$



# 连续 $X$ 之函数



- $Y = g(X)$  且  $X$  为连续随机变数时  
先算  $g(X)$  的CDF :

$$F_{g(X)}(y) = P[g(X) \leq y]$$

- 若  $g(X)$  可微分，再对  $y$  微分得到PDF :

$$f_{g(X)}(y) = \frac{d}{dy} F_{g(X)}(y)$$



# 连续 $X$ 之函数 ( $g(X) = aX + b$ )

- Ex: 若  $Y = 3X + 2$  , 请问  $Y$  的 PDF 跟  $f_X(x)$  之关系为何?



$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(3X + 2 \leq y) \\ &= P\left(X \leq \frac{y-2}{3}\right) \\ &= F_X\left(\frac{y-2}{3}\right) \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y-2}{3}\right) \\ &= \frac{dF_X\left(\frac{y-2}{3}\right)}{d\left(\frac{y-2}{3}\right)} \cdot \frac{d\frac{y-2}{3}}{dy} \\ &= f_X\left(\frac{y-2}{3}\right) \cdot \frac{1}{3} \end{aligned}$$



# 连续 $X$ 之函数 ( $g(X) = aX + b$ )

- 若  $Y = aX + b$  且  $a > 0$ ，则

$$F_Y(y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$



*what if  $a < 0$ ?*

$$F_Y(y) = P(aX + b \leq y) = P\left(\frac{aX + b - b}{a} \geq \frac{y-b}{a}\right) = P\left(X \geq \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right)$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = -\frac{d}{dy} F_X\left(\frac{y-b}{a}\right) = -\frac{dF_X\left(\frac{y-b}{a}\right)}{d\left(\frac{y-b}{a}\right)} \cdot \frac{d\left(\frac{y-b}{a}\right)}{dy} = -f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a} = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$





# 连续 $X$ 之函数 ( $g(X) = aX + b$ )

- Ex: 若  $X \sim \text{Exponential}(\lambda)$ , 且  $Y = 2X$ , 请问  $Y$  之机率分布为何?

$$f_X(x) = \lambda e^{-\lambda x} u(x), Y = 2X \ (a = 2, b = 0)$$

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) = \frac{1}{2} \cdot \lambda e^{-\lambda \frac{y}{2}} \cdot u\left(\frac{y}{2}\right) \\ &= \frac{\lambda}{2} e^{-\frac{\lambda}{2} y} \cdot u(y) \\ &\Rightarrow Y \sim \text{Exponential}\left(\frac{\lambda}{2}\right) \end{aligned}$$



# 连续 $X$ 之函数 ( $g(X) = aX^2 + b$ )

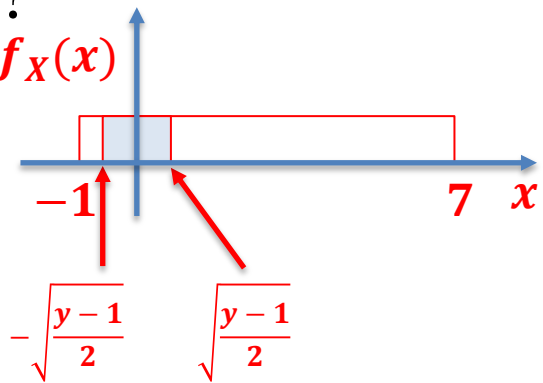
- Ex: 若  $Y = 2X^2 + 1$  , 且已知  $X \sim \text{UNIF}(-1, 7)$  。 请问  $Y$  的 PDF 为何?



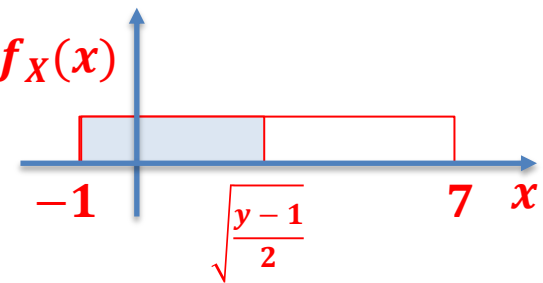
$$\begin{aligned} F_Y(y) &= P(Y = 2X^2 + 1 \leq y) \\ &= P\left(X^2 \leq \frac{y-1}{2}\right) \\ &= P\left(-\sqrt{\frac{y-1}{2}} \leq X \leq \sqrt{\frac{y-1}{2}}\right) \\ &= \int_{-\sqrt{\frac{y-1}{2}}}^{\sqrt{\frac{y-1}{2}}} f_X(x) dx \end{aligned}$$

$y \leq 3: f_X(x)$

$$-1 = -\sqrt{\frac{y-1}{2}} \Rightarrow y = 3$$



$y > 3: f_X(x)$



# 连续 $X$ 之函数 ( $g(X) = aX^2 + b$ )



$$y \leq 3: F_Y(y) = \int_{-\sqrt{\frac{y-1}{2}}}^{\sqrt{\frac{y-1}{2}}} \frac{1}{8} dx = \frac{1}{8} \cdot 2 \cdot \sqrt{\frac{y-1}{2}} = \frac{1}{4} \sqrt{\frac{y-1}{2}}$$

$$y > 3: F_Y(y) = \int_{-1}^{\sqrt{\frac{y-1}{2}}} \frac{1}{8} dx = \frac{1}{8} \left( \sqrt{\frac{y-1}{2}} + 1 \right)$$

$$\Rightarrow y \leq 3: f_Y(y) = \frac{d}{dy} \frac{1}{4} \sqrt{\frac{y-1}{2}} = \frac{1}{4} \cdot \frac{1}{2} \left( \frac{y-1}{2} \right)^{-\frac{1}{2}} \cdot \frac{1}{2} = \frac{1}{16} \sqrt{\frac{2}{y-1}}$$

$$y > 3: f_Y(y) = \frac{d}{dy} \frac{1}{8} \left( \sqrt{\frac{y-1}{2}} + 1 \right) = \frac{1}{32} \sqrt{\frac{2}{y-1}}$$



# 本节回顾

---

- 随机变量的函数又称？
- 若随机变数为离散，可直接推  $g(X)$  之 PMF
- 若随机变数为连续，先推  $g(X)$  之 CDF 再微分得到 PDF 比较好算





# 7-3: 条件机率分布与失忆性

---

第七周



Prof. Yeh, Ping-Cheng (Benson) 葉丙成  
Dept. of EE, National Taiwan University

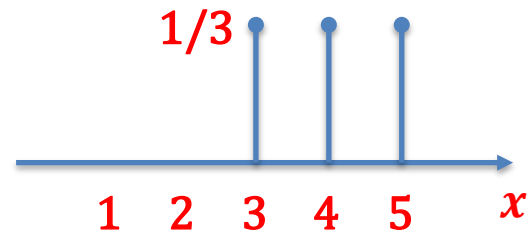
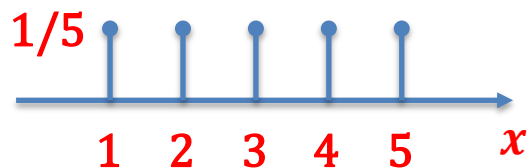
# 把条件机率用在机率分布上



- Ex：为了更了解宅宅的心，店员妹亦开始战 LOL。  
已知店员妹战LOL场数  $X \sim UNIF(1, 5)$ 。若已知  
店员妹战了两场仍战意甚浓、继续战。请问在此情况下，店员妹  
今日战LOL场数  $X$  之机率分布为何？

**B**：已战两场仍想战

$$p_{X|B}(x) = P(X = x|B) = \frac{P(X = x, B)}{P(B)}$$
$$= \begin{cases} \frac{P(X = x)}{P(B)} = \frac{\frac{1}{5}}{\frac{3}{5}} = \frac{1}{3}, & x \in B: x = 3, 4, 5 \\ 0, & x \notin B, x = \text{otherwise} \end{cases}$$



# 条件机率分布 (Conditional Distribution)



- 若  $X$  是一离散随机变数，其 PMF 为  $p_X(x)$ 。若已知某事件  $B$  已发生，则在此情况之下条件机率分布为：

$$\text{– PMF: } p_{X|B}(x) = \begin{cases} x \in B: \frac{p_X(x)}{P(B)}, \\ x \notin B: 0. \end{cases}$$

$$\text{– CDF: } F_{X|B}(x) = \boxed{\sum_{u \leq x} p_{X|B}(u)} = \sum_{u \leq x, u \in B} \frac{p_X(u)}{P(B)}$$



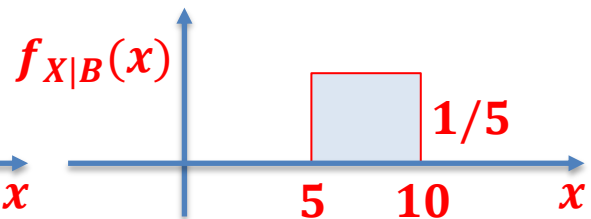
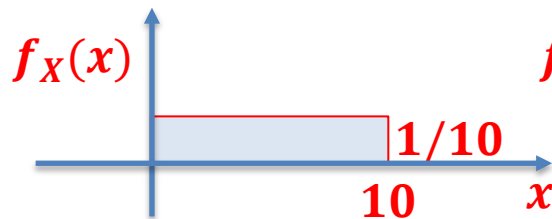
# 把条件机率用在机率分布上



- Ex: 店员妹等公交车上班。通常等公交车的时间  $X$ ，从零到十分钟间可能性均等。若店员妹已等了五分钟车还没来。请问在此情况下，等车时间  $X$  之机率分布为何？

$$X \sim \text{UNIF}(0, 10)$$

$$\boxed{B: X > 5} \Rightarrow P(B) = \frac{1}{2}$$



$$\begin{aligned} f_{X|B}(x) &= \lim_{\Delta \rightarrow 0} \frac{P(X \in [x, x + \Delta] | B)}{\Delta} = \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} \cdot \frac{P(X \in [x, x + \Delta], X \in B)}{P(X \in B)} \\ &= \begin{cases} x \in B: \lim_{\Delta \rightarrow 0} \frac{P(X \in [x, x + \Delta])}{\Delta \cdot P(B)} = \boxed{\frac{f_X(x)}{P(B)}}, \\ x \notin B: 0. \end{cases} \end{aligned}$$





# 条件机率分布 (Conditional Distribution)



- 若  $X$  是一连续随机变量，其 PDF 为  $f_X(x)$ 。若已知某事件  $B$  已发生，则在此情况之下条件机率分布为：

$$\text{– PDF: } f_{X|B}(x) = \begin{cases} x \in B: \frac{f_X(x)}{P(B)}, \\ x \notin B: 0. \end{cases}$$

$$\text{– CDF: } F_{X|B}(x) = \int_{-\infty}^x f_{X|B}(u) du = \int_{-\infty \leq u \leq x, u \in B} \frac{f_X(u)}{P(B)} du$$



# 条件期望值 (Conditional Expectation)

- 若知 **$B$** 已发生，则此情况下条件期望值为：

$$E[X | B] = \begin{cases} \sum_{x=-\infty}^{\infty} x \cdot p_{X|B}(x) & (\text{离散}) \\ \int_{-\infty}^{\infty} x \cdot f_{X|B}(x) dx & (\text{连续}) \end{cases}$$



# 条件期望值 (Conditional Expectation)

- 若知 **$B$** 已发生，则此情况下条件期望值为：



$$E[g(X) | B] = \begin{cases} \sum_{x=-\infty}^{\infty} g(x) \cdot p_{X|B}(x) & \text{(离散)} \\ \int_{-\infty}^{\infty} g(x) \cdot f_{X|B}(x) dx & \text{(连续)} \end{cases}$$



# 条件期望值 (Conditional Expectation)

- 若知 $B$ 已发生，则此情况下条件期望值为：



$$\begin{aligned} \text{Var}(X | B) &= E \left[ (X|_B - \mu_{X|B})^2 \right] = E \left[ (X - \mu_{X|B})^2 \mid B \right] \\ &= E[X^2 | B] - (\mu_{X|B})^2 \end{aligned}$$



# 失忆性 (Memoryless)



- 宅宅与店员妹相约出门。宅宅出门前在战LOL，场数  $X \sim GEO(0.2)$ 。店员妹等了两场后，宅宅还在玩。店员妹甚怒，怒催宅宅。宅宅曰「快好了、快好了」。问宅宅剩余场数  $X'$  之机率分布为何？  $B: X > 2, X' = X|_B - 2$

$$p_{X|B}(x) = \begin{cases} x \in B(x > 2): & \frac{P_X(x)}{P(B)} = \frac{0.8^{x-1} \cdot 0.2}{\sum_{x=3}^{\infty} 0.8^{x-1} \cdot 0.2} = \frac{0.8^{x-1} \cdot 0.2}{0.8^2 \cdot 0.2 \cdot \frac{1}{1-0.8}} = 0.8^{x-3} \cdot 0.2 \\ x \leq 2: & 0 \end{cases}$$

$$p_{X'}(x) = P(X' = x) = P(X|_B - 2 = x) = P(X|_B = x + 2)$$

$$P_{X|B}(x + 2) = 0.8^{x+2-3} \cdot 0.2 = 0.8^{x-1} \cdot 0.2 \Rightarrow X' \sim GEO(0.2)$$



# 失忆性 (Memoryless)



- 店员妹与宅宅相约出门。店员妹出门前化妆时间为  $X$  (小时),  $X \sim \text{Exponential}(1)$ 。经过一小时后, 仍未完成。宅宅甚怒, 怒催店员妹。店员妹曰「快好了、快好了」。问店员妹剩余化妆时间  $X'$  机率分布为何?  $B: X > 1, X' = X|_B - 1$

$$F_{X'}(x) = P(X' \leq x) = P(X|_B - 1 \leq x) = P(X|_B \leq x + 1)$$

$$F_{X|B}(x + 1) = \int_{-\infty}^{x+1} f_{X|B}(u) du = ? \quad F_{X'}(x) = ? \quad f_{X'}(x) = ?$$



# 失忆性 (Memoryless)

$B: X > 1$

$$f_{X|B}(u) = \begin{cases} u \in B: & \frac{f_X(u)}{P(B)} = \frac{1 \cdot e^{-1u}}{P(X > 1)} = \frac{e^{-u}}{1 - F_X(1)} \\ (u > 1) & \\ & = \frac{e^{-u}}{1 - (1 - e^{-1})} = \frac{e^{-u}}{e^{-1}} = e^{-(u-1)} \\ u \notin B: & 0 \end{cases}$$

$$\begin{aligned} \Rightarrow F'_X(x) &= P(X' \leq x) = P(X|_B - 1 \leq x) = P(X|_B \leq x + 1) \\ &= \int_{-\infty}^{x+1} f_{X|B}(u) du = \int_1^{x+1} e^{-(u-1)} du = [-e^{-(u-1)}]_1^{x+1} \\ &= -e^{-x+1-1} - (-e^{-(1-1)}) = 1 - e^{-x} \end{aligned}$$

$$\Rightarrow f_{X'}(x) = 1 \cdot e^{-1 \cdot x}, x \geq 0; 0 \text{ otherwise.} \Rightarrow X' \sim \text{Exponential}(1)$$



# 失忆性 (Memoryless)

---

- Geometric 跟 Exponential 机率分布皆有失忆性的性质
- 不管事情已经进行多久，对于事情之后的进行一点影响都没有！





# 本节回顾

- 某个事件发生后，随机变量的行为跟其机率分布也会改变：条件机率分布
- 条件随机变量也是一个健全、可爱的随机变数！
- 身为一个健全、可爱的随机变数，人家一般随机变数该有的条件随机变数也应该都有！PMF (PDF)、CDF、期望值、Mean、Variance 等
- 随机变数中会失忆的是？

