Probability & Statistics (1)

Continuous Random Variables

Asst. Prof. Chan, Chun-Hsiang

Master program in Intelligent Computing and Big Data, Chung Yuan Christian University, Taoyuan, Taiwan
Undergraduate program in Intelligent Computing and Big Data, Chung Yuan Christian University, Taoyuan, Taiwan
Undergraduate program in Applied Artificial Intelligence, Chung Yuan Christian University, Taoyuan, Taiwan

Outlines

1. Introduction
2. Expectation and Variance of Continuous Random Variables
3. The Uniform Random Variable
4. Normal Random Variables
5. Exponential Random Variables
6. The Distribution of a Function of a Random Variable
7. [#10] Assignment
8. Reference
9. Question Time

Introduction

- 上一章節都是在講離散隨機變數，本章節主要介紹連續隨機變數，所以我們可以討論連續空間上的機率性質，譬如說: 生命長短、到達時間等。
- 如果我們給定一個 X 為連續隨機變數 (continuous random variable)，存在一個非負函數(nonnegative function) f，定義對於所有實數 $x \in \{-\infty, \infty\}$，則對於 B 集合中的所有實數，

$$P\{X \in B\} = \int_B f(x)dx$$

- f 函數為隨機變數 X 的機率密度函數(probability density function)。

Introduction

- Since X must assume some value, f must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

- $let\ B = [a, b], then$

$$P\{a \le X \le b\} = \int_{a}^{b} f(x)dx$$

- $If\ we\ let\ a = b, then$

$$P\{X = a\} = \int_{a}^{a} f(x)dx = 0$$

- For a continuous random variable, we can define its cumulative density function (c.d.f.) as follows,

$$P\{X < a\} = P\{X \le a\} = F(a) = \int_{-\infty}^{a} f(x)dx$$

Introduction

- **範例一**

令X為continuous random variable，其PDF被定義如下:

$$f(x) = \begin{cases} C(4x - 2x^2), 0 < x < 2 \\ 0, otherwise \end{cases}$$

(1) $C = ?$

(2) $P\{X > 1\} = ?$

Introduction

Solution:

(1) Since f is a probability density function, we must have $\int_{-\infty}^{\infty} f(x)dx = 1$, implying that

$$C \int_0^2 (4x - 2x^2)dx = 1$$

$$C \left[2x^2 - \frac{2}{3}x^3\right] \Big|_{x=0}^{x=2} = 1$$

$$C = \frac{3}{8}$$

(2) $P\{X > 1\} = \int_1^{\infty} f(x)dx = \frac{3}{8}\int_1^2 (4x - 2x^2)dx = \frac{1}{2}$

Introduction

- **範例二**

如果搭機前往卡達看世足的飛行時間為一個continuous random variable，其PDF可以被定義為

$$f(x) = \begin{cases} \lambda e^{-\frac{x}{100}}, x \geq 0 \\ 0, x < 0 \end{cases}$$

試問以下的機率:

(1) 飛行時間介於50-150分鐘的機率?

(2) 飛行時間小於100分鐘的機率?

Introduction

Solution:

(1) Since

$$1 = \int_{-\infty}^{\infty} f(x)dx = \lambda \int_{0}^{\infty} e^{-\frac{x}{100}} dx$$

$$1 = -\lambda(100)e^{-\frac{x}{100}}\Big|_{0}^{\infty} = 100\lambda \Rightarrow \lambda = \frac{1}{100}$$

$Hence,$

$$P\{50 < X < 150\} = \int_{50}^{150} \frac{1}{100} e^{-\frac{x}{100}} dx = -e^{-\frac{x}{100}}\Big|_{50}^{150} = -e^{-\frac{2}{3}} - (-e^{-\frac{1}{2}})$$

$$\approx 0.384$$

Introduction

(2)

$$P\{X < 100\} = \int_0^{100} \frac{1}{100} e^{-\frac{x}{100}} dx = -e^{-\frac{x}{100}} \Big|_0^{100} = 1 - e^{-1} \approx 0.633$$

[加分題]

(2) 0.633是甚麼意思呢?

Introduction

- **範例三**

如果一根紫外線燈管的壽命為一個continuous random variable，其PDF可以被定義為

$$f(x) = \begin{cases} 0, x \leq 100 \\ \dfrac{100}{x^2}, x > 100 \end{cases}$$

試問五個之中有兩個燈管在開始使用150分鐘後就壞掉的機率?

Solution:

假設事件$E_i, i = 1,2,3,4,5$，第$i - th$燈管壞掉的事件彼此獨立。

Introduction

$$P(E_i) = \int_0^{150} f(x)dx = 100\int_{100}^{150} x^{-2}dx = \frac{1}{3}$$

Hence,

$$P\left(\frac{2}{5} were\ broken\ in\ the\ first\ 150\ \text{min}\right) = \binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 = \frac{80}{143}$$

Introduction

$Cumulative\ distribution\ F\ could\ be\ obtained\ from\ the\ p.d.f.$

$$F(a) = P\{X \in (-\infty, a]\} = \int_{-\infty}^{a} f(x)dx$$

$Differentiating\ both\ side$

$$\frac{d}{da}F(a) = f(a)$$

$$\because \int_{a}^{a} f(x)dx = 0$$

$$\therefore P\left\{a - \frac{\varepsilon}{2} \le X \le a + \frac{\varepsilon}{2}\right\} = \int_{a-\varepsilon/2}^{a+\varepsilon/2} f(x)dx \approx \varepsilon f(a), where\ \varepsilon\ is\ very\ small.$$

Introduction

- **範例四**

如果X為continuous with distribution function F_X 與 density function f_X，試求$Y = 2X$的distribution function。

Solution 1:

$$F_Y(a) = P\{Y \leq a\} = P\{2X \leq a\} = P\left\{X \leq \frac{a}{2}\right\} = F_X\left(\frac{a}{2}\right)$$

$$differentiation$$

$$f_Y(a) = \frac{1}{2} f_X\left(\frac{a}{2}\right)$$

Introduction

Solution 2:

$$\epsilon f_Y(a) \approx P\left\{a - \frac{\epsilon}{2} \le Y \le a + \frac{\epsilon}{2}\right\}$$

$$= P\left\{a - \frac{\epsilon}{2} \le 2X \le a + \frac{\epsilon}{2}\right\}$$

$$= P\left\{\frac{a}{2} - \frac{\epsilon}{4} \le X \le \frac{a}{2} + \frac{\epsilon}{4}\right\} \approx \frac{\epsilon}{2} f_X\left(\frac{a}{2}\right)$$

$dividing\ by\ \epsilon$

$$f_Y(a) \approx \frac{1}{2} f_X\left(\frac{a}{2}\right)$$

Expectation and Variance of Continuous Random Variables

- 在discrete random variable計算expected value時:

$$E[X] = \sum_x xP\{X = x\}$$

- 如果X為continuous random variable with $p.d.f.\ f(x)$，則:

$$f(x)dx \approx P\{x \le X \le x + dx\}\ for\ dx\ small$$

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

Expectation and Variance of Continuous Random Variables

- 範例五

試求出$E[X]$，當X的density function為:

$$f(x) = \begin{cases} 2x, if\ 0 \le x \le 1 \\ 0, otherwise \end{cases}$$

Solution:

$$E[X] = \int xf(x)dx = \int_0^1 2x^2 dx = \frac{2}{3}$$

Expectation and Variance of Continuous Random Variables

- 範例六

當X的density function為:

$$f(x) = \begin{cases} 1, if\ 0 \leq x \leq 1 \\ 0, otherwise \end{cases}$$

試求$E[e^X]$:

Solution:

Let $Y = e^X$. Now we need to determine F_Y, and the probability distribution function of Y, where $1 \leq x \leq e$.

Expectation and Variance of Continuous Random Variables

$$F_Y = P\{Y \le x\} = P\{e^x \le x\} = P\{X \le \log x\} = \int_0^{\log(x)} f(y)dy = logx$$

By differentiating $F_Y(x)$, the $p.d.f.$ of Y is given by…

$$f_Y(x) = \frac{1}{x}, where\ 1 \le x \le e$$

Hence,

$$E[e^X] = E[Y] = \int_{-\infty}^{\infty} x f_Y(x)dx = \int_1^e dx = e - 1$$

Expectation and Variance of Continuous Random Variables

- **Proposition 1**

If X is a continuous random variable with probability density function $f(x)$, then. For any real-valued function g,

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

$example$:

$$E[e^X] = \int_0^1 e^x \, dx = e - 1, since\ f(x) = 1, where\ 0 < x < 1$$

Expectation and Variance of Continuous Random Variables

- **Lemma 1**

For nonnegative random variable Y,

$$E[Y] = \int_0^\infty P\{Y > y\}dy$$

Proof:

We present a proof when Y is a continuous random variable with probability density function f_Y. We have

$$\int_0^\infty P\{Y > y\}\,dy = \int_0^\infty \int_y^\infty f_Y(x)dxdy$$

Expectation and Variance of Continuous Random Variables

Where we have used the fact that $P\{Y > y\} = \int_y^{\infty} f_Y(x)dx$.

Interchanging the order of integration in the proceeding equation yields.

$$\int_0^{\infty} P\{Y > y\}dy = \int_0^{\infty} \left(\int_0^{x} dy \right) f_Y(x)dx = \int_0^{\infty} x f_Y(x)dx = E[Y]$$

Expectation and Variance of Continuous Random Variables

Proof of Proposition 1:

From **Lemma 1**, for any function g for which $g(x) \geq 0$.

$$E[g(X)] = \int_0^\infty P\{g(X) > y\}dy = \int_0^\infty \int_{x:g(x)>y} f(x)dxdy$$

$$= \int_{x:g(x)>0} \int_0^{g(x)} dyf(x)dx = \int_{x:g(x)>0} g(x)f(x)dx$$

Expectation and Variance of Continuous Random Variables

- 範例七

如果今天有一根長度為1的竹筷，你在點U的地方折斷，且U符合uniformly distributed，值域落在$(0,1)$。試問期望的長度中含有點p的機率為何 $(0 \leq p \leq 1)$?

Solution:

令$L_p(U)$為折斷後的竹筷含有點p的長度:

$$L_p(U) = \begin{cases} 1 - U, U < p \\ U, U > p \end{cases}$$

Expectation and Variance of Continuous Random Variables

From **Proposition 1**,

$$E\left[L_p(U)\right] = \int_0^1 L_p(u)du = \int_0^p (1-u)du + \int_p^1 udu$$

$$= \frac{1}{2} - \frac{(1-p)^2}{2} + \frac{1}{2} - \frac{p^2}{2} = \frac{1}{2} + p(1-p)$$

Expectation and Variance of Continuous Random Variables

- **範例八**

假設你提早s分鐘球場練球，你所需要花的成本為cs；如果你晚s分鐘到球場，你的成本則為ks。假設你從你的所在地出發前往到球場的時間為continuous random variable，且$p.d.f.$為f。試問你應該要幾點出發使得你得成本最小化。

Solution:

令X為旅行時間。如果你提早t分鐘離開，則你的成本為$C_t(X)$:

$$C_t(X) = \begin{cases} c(t-X), if\ X \le t \\ k(X-t), if\ X \ge t \end{cases}$$

Expectation and Variance of Continuous Random Variables

因此

$$E[C_t(X)] = \int_0^\infty C_t(x)f(x)dx = \int_0^t c(t-x)f(x)dx + \int_t^\infty k(x-t)f(x)dx$$

$$= ct\int_0^t f(x)dx - c\int_0^t xf(x)dx + k\int_t^\infty xf(x)dx - kt\int_t^\infty f(x)dx$$

我們要做的就是最小化$E[C_t(X)]$，利用微分可以得出

$$\frac{d}{dt}E[C_t(X)] = ctf(t) + cF(t) - ctf(t) - ktf(t) + ktf(t) - k[1-F(t)]$$

$$= (k+c)F(t) - k$$

Expectation and Variance of Continuous Random Variables

令$(k+c)F(t)-k$等於0

則可求出最佳化的t^*時間為

$$F(t^*)=\frac{k}{k+c}$$

Corollary 1

If a and b are constant, then

$$E[aX+b]=aE[X]+b$$
$$Var(X)=E[(X-\mu)^2]=E[X^2]-(E[X])^2$$

Expectation and Variance of Continuous Random Variables

- **範例九**

試求出$Var(X)$，當X的$p.d.f.$為

$$f(x) = \begin{cases} 2x, if\ 0 \le x \le 1 \\ 0, otherwise \end{cases}$$

Solution:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 2x^3 dx = \frac{1}{2}$$

$$since\ E[X] = \frac{2}{3}$$

$$Var(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

The Uniform Random Variable

如果說一個continuous random variable是uniformly distributed分布在$(0,1)$的區間中，則其$p.d.f.$為

$$f(x) = \begin{cases} 1, 0 < x < 1 \\ 0, otherwise \end{cases}$$

Since $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = \int_0^1 dx = 1$. Because $f(x) > 0$ only when $x \in (0,1)$, it follows that X must assume a value in interval $(0,1)$. Also, since $f(x)$ is constant for $x \in (0,1)$, X is just as likely to be near any value in $(0,1)$ as it is to be near any other value.

To verify this statement, note that, for any $0 < a < b < 1$,

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx = b - a$$

The Uniform Random Variable

$$P\{a \le X \le b\} = \int_a^b f(x)dx = b - a$$

We say that X is a uniform random variable on the interval (α, β) if the probability density function of X is given by,

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, if\ \alpha < x < \beta \\ 0, otherwise \end{cases}$$

Since $F(a) = \int_{-\infty}^{a} f(x)dx$, then the cumulative density function is given by,

$$F(a) = \begin{cases} 0, a \le \alpha \\ \frac{a - \alpha}{\beta - \alpha}, \alpha < x < \beta \\ 1, a \ge \beta \end{cases}$$

The Uniform Random Variable

- 範例十

令X為uniformly distributed over (α, β)。 試求(a) $E[X]$; (b) $Var(X)$

Solution:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{\alpha}^{\beta} \frac{x}{\beta - \alpha} dx = \frac{\beta^2 - \alpha^2}{2(\beta - \alpha)} = \frac{\beta + \alpha}{2}$$

$$E[X^2] = \int_{\alpha}^{\beta} \frac{1}{\beta - \alpha} x^2 dx = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha\beta + \alpha^2}{3}$$

$$Var(X) = \frac{\beta^2 + \alpha\beta + \alpha^2}{3} - \left(\frac{\beta + \alpha}{2}\right)^2 = \frac{(\beta - \alpha)^2}{12}$$

The Uniform Random Variable

- **範例十一**

如果X是uniformly distributed且值域落在(0,10)，請計算下列機率:

(a) $X < 3$; (b) $X > 6$; (c) $3 < X < 8$.

Solution:

(a)

$$P\{X < 3\} = \int_0^3 \frac{1}{10}\,dx = \frac{3}{10}$$

(b)

$$P\{X > 6\} = \int_6^{10} \frac{1}{10}\,dx = \frac{4}{10}$$

(c)

$$P\{3 < X < 8\} = \int_3^8 \frac{1}{10}\,dx = \frac{1}{2}$$

The Uniform Random Variable

- **範例十二**

假設往返學校與火車站的接駁車從早上七點開始，每十五分鐘一班車，也就是說靠站時間為: 7:00, 7:15, 7:30, 7:45, …。試問一名乘客等車的時間為uniformly distributed，且時間落在於7:00與7:30之間，試問他等車的機率:

(a) 等候時間小於五分鐘

(b) 整候時間大於十分鐘

The Uniform Random Variable

Solution:

令X為七點後等車的時間，X為uniform random variable且值域落在(0,30)之間，如果等車時間要小於5分鐘的話，則只有兩種:

7:10-7:15 or 7:25-7:30

$$P\{10 < X < 15\} + P\{25 < X < 30\} = \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}$$

同理，如果等車時間會超過十分鐘的話:

7:00-7:05 or 7:15-7:20

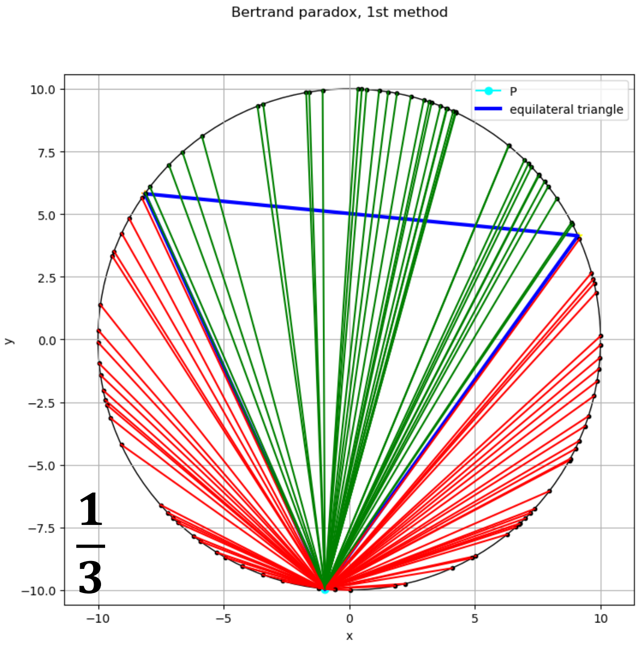
$$P\{0 < X < 5\} + P\{15 < X < 20\} = \frac{1}{3}$$

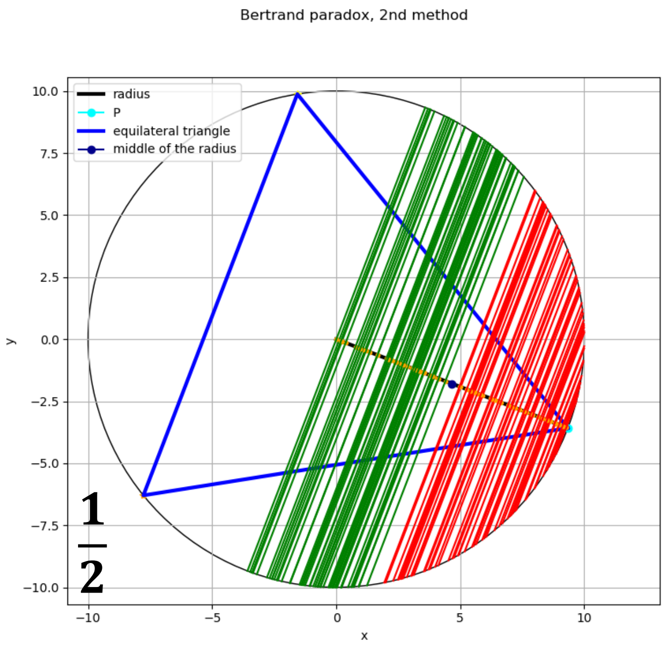
The Uniform Random Variable
Bertrand's Paradox

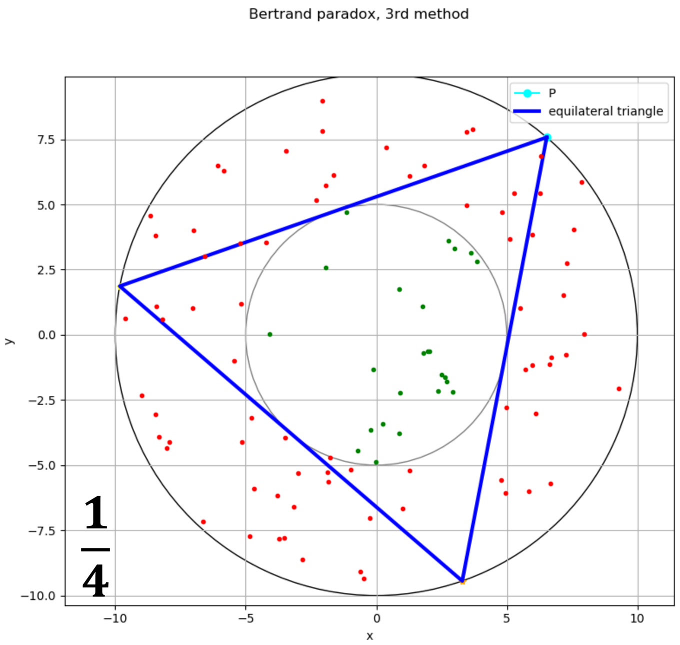
https://github.com/czimbortibor/Bertrand-paradox/blob/master/bertrand.py

為甚麼我們要那麼認真講機率的定義? 如果沒有好好的定義會有差?

Consider a random chord (弦) of a circle. What is the probability that the length of the chord will be greater than the side of the equilateral triangle inscribed in that circle?



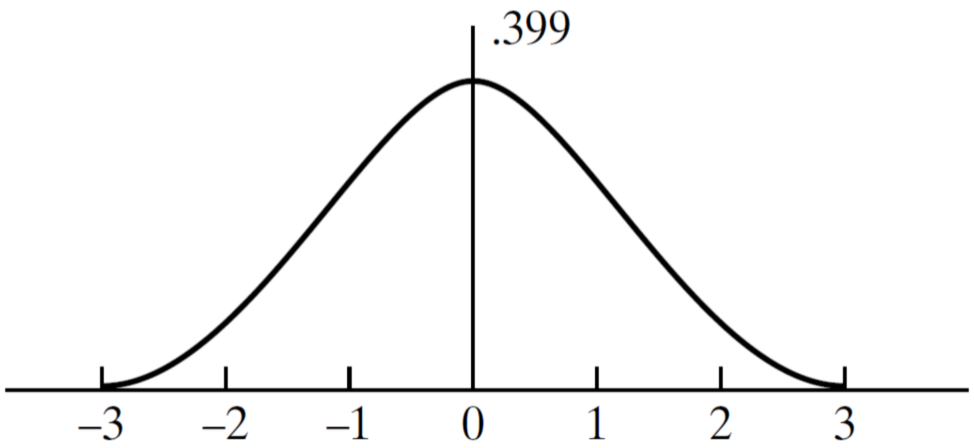


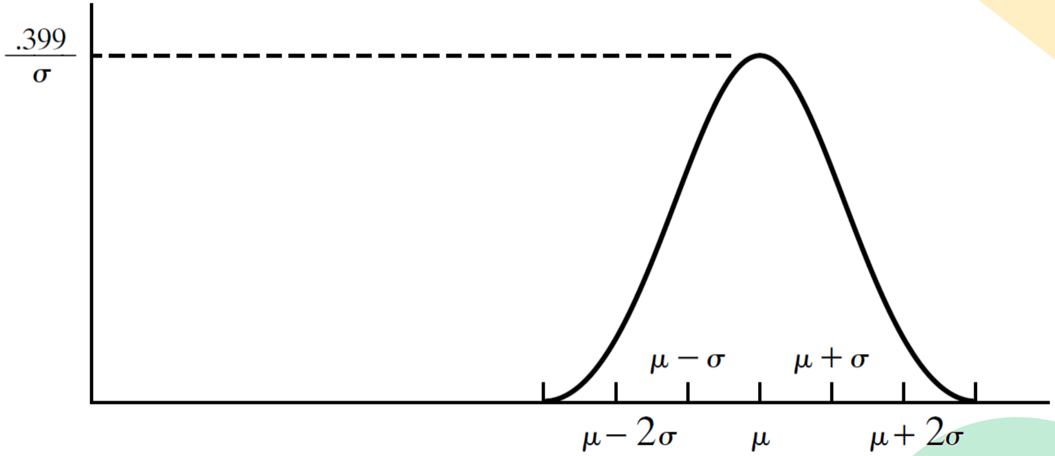


Normal Random Variables

- 如果X為一個normal random variable，或者是說X是normally distributed，具有兩個參數(μ,σ^2)，如果X的$p.d.f.$為:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, where\ -\infty<x<\infty$$





Normal Random Variables

To prove that $f(x)$ is indeed a probability density function, we need to show that

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = 1$$

Let $y = \frac{x-\mu}{\sigma}$

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy = 1$$

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy = \sqrt{2\pi}$$

Normal Random Variables

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy = \sqrt{2\pi}$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2}\, dy$

$$I^2 = \int_{-\infty}^{\infty} e^{-y^2/2}\, dy \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{y^2+x^2}{2}}\, dydx$$

Let $x = rcos\theta; y = rsin\theta; dydx = rd\theta dr$.

$$I^2 = \int_0^{\infty}\int_0^{2\pi} e^{-\frac{r^2}{2}}\, rd\theta dr = 2\pi \int_0^{\infty} re^{-\frac{r^2}{2}}\, dr = -2\pi e^{-\frac{r^2}{2}}\Big|_0^{\infty} = 2\pi$$

$$\Rightarrow I = \sqrt{2\pi}$$

Normal Random Variables

If X is normally distributed ➔ $X \sim normal(\mu, \sigma^2) \Rightarrow Y = aX + b \Rightarrow Y \sim normal(a\mu + b, a^2\sigma^2)$

Let F_Y denote the cumulative distribution function of Y.

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

Where F_X is the $c.d.f.$ of X. By differentiation, the density function of Y is…

$$f_Y = \frac{1}{a} f_X\left(\frac{x-b}{a}\right) = \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{\left(\frac{x-b}{a}-\mu\right)^2}{2\sigma^2}} = \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{(x-b-a\mu)^2}{2a^2\sigma^2}}$$

Normal Random Variables

- **範例十三**

試求出$E[X]$與$Var(X)$，當X為normal random variable，且其參數為μ與σ^2。

Solution:

令我們找的mean and variance都來自standard normal random variable $Z = (X - \mu)/\sigma$。

$$E[Z] = \int_{-\infty}^{\infty} x f_Z(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} = 0$$

$$Var(Z) = E[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx$$

Normal Random Variables

$$Var(Z) = E[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx$$

Integration by parts (with $u = x$ and $dv = xe^{-\frac{x^2}{2}}$), …

$$Var(Z) = \frac{1}{\sqrt{2\pi}} \left(-xe^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right)$$

$$= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \right) = 1$$

Normal Random Variables

Because $X = \mu + \sigma Z$, …

$$E[X] = \mu + \sigma E[Z] = \mu$$
$$Var(X) = \sigma^2 Var(Z) = \sigma^2$$

c.d.f. of a standard normal random variable by $\Phi(x)$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}}\, dy$$

Normal Random Variables

The value of $\Phi(x)$ for nonnegative x are given

$$\Phi(-x) = 1 - \Phi(x), -\infty < x < \infty$$

If Z is a standard normal random variable, then

$$P\{Z \leq -x\} = P\{Z > x\}$$

Since $Z = \frac{X-\mu}{\sigma}$

$$F_X(a) = P\{X \leq a\} = P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{a-\mu}{\sigma}\right)$$

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Normal Random Variables

- **範例十四**

如果$X \sim normal(\mu = 3, \sigma^2 = 9)$，試問:

(a) $P\{2 < X < 5\}$

(b) $P\{X > 0\}$

(c) $P\{|X - 3| > 6\}$

Solution:

(a)

$$P\{2 < X < 5\} = P\left\{\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right\} = P\left\{-\frac{1}{3} < Z < \frac{2}{3}\right\}$$

$$= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) = \Phi\left(\frac{2}{3}\right) - \left[1 - \Phi\left(\frac{1}{3}\right)\right] \approx 0.3779$$

Normal Random Variables

(b)

$$P\{X > 0\} = P\left\{\frac{X-3}{3} > \frac{0-3}{3}\right\} = P\{Z > -1\} = 1 - \Phi(-1) = \Phi(1)$$

$$\approx 0.8413$$

(c)

$$P\{|X-3| > 6\} = P\{X > 9\} + P\{X < -3\}$$

$$= P\left\{\frac{X-3}{3} > \frac{9-3}{3}\right\} + P\left\{\frac{X-3}{3} < \frac{-3-3}{3}\right\} = P\{Z > 2\} + P\{Z < -2\}$$

$$= 1 - \Phi(2) + \Phi(-2) = 2[1 - \Phi(2)] \approx 0.0456$$

Normal Random Variables

- **範例十五**

假設今天全班的期末考成績為鐘型曲線(normal distribution)，令學生的成績為$X \sim normal(\mu, \sigma^2)$，老師打算依照成績分布給等第:

(a) A成績大於$\mu + \sigma$

(b) B成績介於μ與$\mu + \sigma$之間

(c) C成績介於$\mu - \sigma$與μ之間

(d) D成績介於$\mu - \sigma$與$\mu - 2\sigma$

(e) F成績小於$\mu - 2\sigma$

Normal Random Variables

Solution:

$$P\{X > \mu + \sigma\} = P\left\{\frac{X-\mu}{\sigma} > 1\right\} = 1 - \Phi(1) \approx 0.1587$$

$$P\{\mu < X < \mu + \sigma\} = P\left\{0 < \frac{X-\mu}{\sigma} < 1\right\} = \Phi(1) - \Phi(0) \approx 0.3413$$

$$P\{\mu - \sigma < X < \mu\} = P\left\{-1 < \frac{X-\mu}{\sigma} < 0\right\} = \Phi(0) - \Phi(-1) \approx 0.3413$$

$$P\{\mu - 2\sigma < X < \mu - \sigma\} = P\left\{-2 < \frac{X-\mu}{\sigma} < -1\right\} = \Phi(2) - \Phi(1) \approx 0.1359$$

$$P\{X < \mu - 2\sigma\} = P\left\{\frac{X-\mu}{\sigma} < -2\right\} = \Phi(-2) \approx 0.0228$$

Normal Random Variables

- **範例十六**

假設X為小孩出生的時間(一年中date為單位)，所以可以得到$X \sim normal(\mu, = 270\ , \sigma^2 = 10)$。試問: 出生時間晚於290天或早於240天的機率為何?

Solution:

$$P\{X > 290\ or\ X < 240\} = P\{X > 290\} + P\{X < 240\}$$

$$= P\left\{\frac{X-270}{10} > 2\right\} + P\left\{\frac{X-270}{10} > -3\right\} = 1 - \Phi(2) + 1 - \Phi(3)$$

$$\approx 0.0241$$

Normal Random Variables

- **範例十七**

假設要從A地傳遞一個二元的資料(0,1)到B地，為了讓訊號辨別得更好，2代表1，而−2代表0。因此x值域為$x = \pm 2$。收到的訊號為R，其中還會接受到一些雜訊N，也就是$R = x + N$。訊號拆解方式如下:

$$If\ R \geq 0.5, then\ 1\ is\ concluded.$$
$$If\ R < 0.5, then\ 0\ is\ concluded.$$

$$P\{error \mid message\ is\ 1\} = P\{M < -1.5\} = 1 - \Phi(1.5) \approx 0.0668$$
$$P\{error \mid message\ is\ 0\} = P\{M \geq 2.5\} = 1 - \Phi(2.5) \approx 0.0062$$

Normal Random Variables
The Demoivre-Laplace Limit Theroem

- If S_n denotes the number of success that occur when n independent trials, each resulting in a success with probability p, are perform, then, for $a < b$.

$$P\left\{a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right\} \rightarrow \Phi(b) - \Phi(a), as\, n \rightarrow \infty$$

Normal Random Variables
The Demoivre-Laplace Limit Theroem

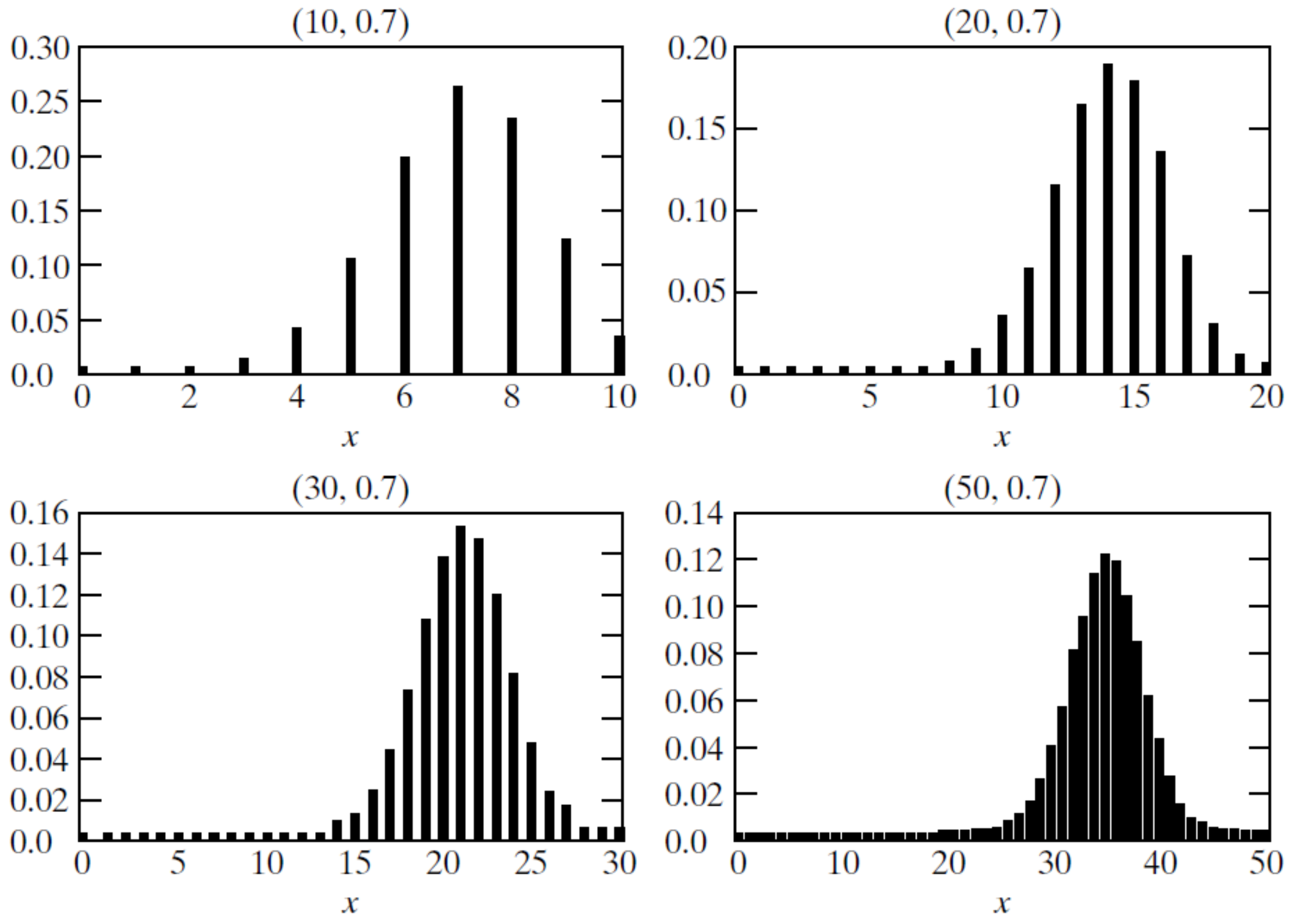


FIGURE 5.6: The probability mass function of a binomial (n, p) random variable becomes more and more "normal" as n becomes larger and larger.

Normal Random Variables

- **範例十八**

令X為投擲一枚公平的硬幣40次中正面的次數，試問$X = 20$的機率

Solution:

$$P\{X = i\} = P\{19.5 \le X \le 20.5\}$$
$$= P\left\{\frac{19.5 - 20}{\sqrt{10}} < \frac{X - 20}{\sqrt{10}} < \frac{20.5 - 20}{\sqrt{10}}\right\}$$
$$\approx P\left\{-0.16 < \frac{X - 20}{\sqrt{10}} < 0.16\right\}$$
$$\approx \Phi(0.16) - \Phi(-0.16) \approx 0.1272$$

Normal Random Variables

- **範例十九**

假設現在進行系上招生，理想班級人數為150人，但只有30%人會入學，如果有450人報名，試問等於或超過150人入學的機率為何?

Solution:

Let X is a binomial random variable, where parameters $n = 450$ and $p = 0.3$。考慮到連續空間，我們就可以用常態來做逼近求值。

$$P\{X >= 150.5\} = P\left\{\frac{X - 450 \times 0.3}{\sqrt{450(0.3)(0.7)}} \geq \frac{150 - 450 \times 0.3}{\sqrt{450(0.3)(0.7)}}\right\}$$

$$\approx 1 - \Phi(1.59) \approx 0.0559$$

Normal Random Variables

- 範例二十

假設現在有一款新型高血壓藥物，有100人服用此藥物，至少有65%人可以成功降低血壓，則現在有一個新病患要服用此藥物，試問降低血壓的機率為何?

Solution:

令X為降低血壓的人數為常態分佈，且一個人是否降低血壓機率為$p = 1/2$。

$$\sum_{i=65}^{100} \binom{100}{i} \left(\frac{1}{2}\right)^{100} = P\{X \geq 64.5\} = P\left\{\frac{X - 100\left(\frac{1}{2}\right)}{\sqrt{100\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}} \geq 2.9\right\}$$

$$\approx 1 - \Phi(2.9) \approx 0.0019$$

Exponential Random Variables

- 若X為Exponential random variable，其$p.d.f.$為

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, if x \geq 0 \\ 0, if\ x < 0 \end{cases}$$

- 其$c.d.f.$為

$$F(a) = P\{X \leq a\} = \int_0^a \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_0^a = 1 - e^{-\lambda a}, where\ a > 0$$

Note that $F(\infty) = \int_0^\infty \lambda e^{-\lambda x} dx = 1$, as, of course, it must. The parameter λ will now be shown to equal the reciprocal of the expected value.

Exponential Random Variables

- 範例二十一

令X為exponential random variable參數為λ。

試問: (a) $E[X]$; (b) $Var(X)$.

Solution:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, if x \geq 0 \\ 0, if\ x < 0 \end{cases}$$

For $n > 0$,

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx$$

Exponential Random Variables

$$E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx$$

Integration by parts (with $\lambda e^{-\lambda x} = dv$ and $u = x^n$) yields

$$E[X^n] = -x^n e^{-\lambda x} \Big|_0^\infty + \int_0^\infty e^{-\lambda x} n x^{n-1} dx$$

$$= 0 + \frac{n}{\lambda} \int_0^\infty x^{n-1} \lambda e^{-\lambda x} dx$$

$$= \frac{n}{\lambda} E[X^{n-1}]$$

Exponential Random Variables

$$\frac{n}{\lambda} E[X^{n-1}]$$

Let $n = 1$ and $n = 2$

$$E[X] = \frac{1}{\lambda}; E[X^2] = \frac{2}{\lambda} E[X] = \frac{2}{\lambda^2}$$

(b)

$$Var(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Exponential Random Variables

- **範例二十二**

假設X為搭火車的時間$X\sim exponential(\lambda = \frac{1}{10})$，今天你去搭火車從某一站到另一站:

(a) 旅行時間大於10分鐘的機率

(b) 旅行時間介於10與20分鐘之間的機率

Solution:

(a) $P\{X > 10\} = 1 - F(10) = e^{-1} \approx 0.368$

(b) $P\{10 < X < 20\} = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.233$

Exponential Random Variables

We say that a nonnegative random variable X is memoryless if

$$P\{X > s + t \mid X > t\} = P\{X > s\}, \qquad for\ all\ s, t \geq 0$$

It is equivalent to …

$$\frac{P\{X > s + t, X > t\}}{P\{X > t\}} = P\{X > s\}$$

$$P\{X > s + t, X > t\} = P\{X > s\}P\{X > t\}$$

The Distribution of a Function of a Random Variable

- 範例二十三

令$X\sim uniform(0,1)$，試問隨機變數$Y=X^n$ (For $0\leq y\leq 1$)的$p.d.f.$。

$$
\begin{aligned}
F_Y(y) &= P\{Y\leq y\} \\
&= P\{X^n\leq y\} \\
&= P\left\{X\leq y^{\frac{1}{n}}\right\} \\
&= F_X\left(y^{\frac{1}{n}}\right)
\end{aligned}
$$

Differentiation to obtain the $p.d.f.$

$$
f_Y(y)=\begin{cases}\frac{1}{n}y^{\frac{1}{n}-1}, 0\leq y\leq 1 \\ 0, otherwise\end{cases}
$$

The Distribution of a Function of a Random Variable

- 範例二十四

令X為continuous random variable，$p.d.f.$為f_X，distribution為 $Y = X^2$ $(For\ y \geq 0)$:

$$
\begin{aligned}
F_Y &= P\{Y \leq y\} \\
&= P\{X^2 \leq y\} \\
&= P\{-\sqrt{y} < X < \sqrt{y}\} \\
&= F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}
$$

Differentiation yields

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) - f_X(-\sqrt{y})]$$

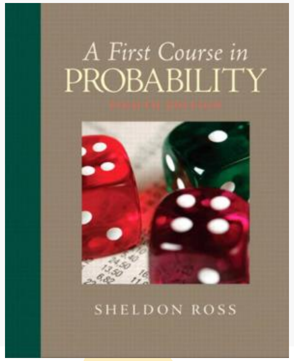
The Distribution of a Function of a Random Variable

[加分題]

令X為continuous random variable，$p.d.f.$為f_X，distribution為
$Y = |X|\ (For\ y >= 0)$:

…

[#10] Assignment



- Selected Problems from Sheldon Ross Textbook [1].

5.15. If X is a normal random variable with parameters $\mu = 10$ and $\sigma^2 = 36$, compute
(a) $P\{X > 5\}$;
(b) $P\{4 < X < 16\}$;
(c) $P\{X < 8\}$;
(d) $P\{X < 20\}$;
(e) $P\{X > 16\}$.

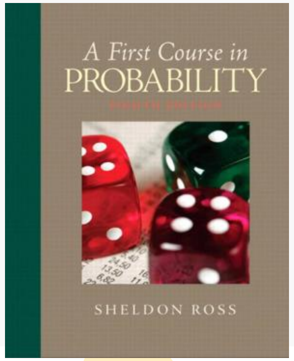
5.16. The annual rainfall (in inches) in a certain region is normally distributed with $\mu = 40$ and $\sigma = 4$. What is the probability that, starting with this year, it will take over 10 years before a year occurs having a rainfall of over 50 inches? What assumptions are you making?

5.17. A man aiming at a target receives 10 points if his shot is within 1 inch of the target, 5 points if it is between 1 and 3 inches of the target, and 3 points if it is between 3 and 5 inches of the target. Find the expected number of points scored if the distance from the shot to the target is uniformly distributed between 0 and 10.

5.18. Suppose that X is a normal random variable with mean 5. If $P\{X > 9\} = .2$, approximately what is $\text{Var}(X)$?

[1] Sheldon Ross. A First of Course in Probability. 8th edition.

[#10] Assignment



5.19. Let X be a normal random variable with mean 12 and variance 4. Find the value of c such that $P\{X > c\} = .10$.

5.20. If 65 percent of the population of a large community is in favor of a proposed rise in school taxes, approximate the probability that a random sample of 100 people will contain

(a) at least 50 who are in favor of the proposition;

(b) between 60 and 70 inclusive who are in favor;

(c) fewer than 75 in favor.

5.7. The standard deviation of X, denoted $SD(X)$, is given by

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

Find $SD(aX + b)$ if X has variance σ^2.

[1] Sheldon Ross. A First of Course in Probability. 8th edition.

Reference

Ross, S. (2010). *A first course in probability*. Pearson.

The End

If you have any questions, please do not hesitate to ask me.

Thank you for your attention))