

Probability_HW07_Solution

Problem 16

(★) : Assume that the events of annual rainfall exceeding 50 inches each year are mutually independent.

Let X be the annual rainfall for a given year. Then $X \sim N(40, 4^2)$.

Let $p \equiv \Pr(X < 50)$. Then we have :

$$\begin{aligned} p &= \Pr\left(\frac{X-40}{4} < \frac{50-40}{4}\right) \\ &= \Pr(Z < 2.5) \quad \left(\text{By normalizing } X, \text{ we have } Z \equiv \frac{X-40}{4} \sim N(0, 1).\right) \\ &= \Phi(2.5) = 0.9937903 \end{aligned}$$

Let Y be the number of years starting from this year until the first occurrence of annual rainfall exceeding 50 inches. Then, $Y \sim \text{Geometric}(1-p)$, and the desired quantity is :

$$\begin{aligned} &\Pr(\text{None of the following 10 years has a rainfall of more than 50 inches}) \\ &\stackrel{(*)}{=} \Pr(Y > 10) = \sum_{y=11}^{\infty} p^{y-1}(1-p) = p^{10} = (0.9937903)^{10} \end{aligned}$$

```
pnorm(2.5)
```

```
## [1] 0.9937903
```

Problem 17

Let X be the salaries. Then, $X \sim N(\mu, \sigma^2)$, and we have :

$$\begin{aligned} \begin{cases} 0.25 = \Pr(X < 180,000) \\ 0.25 = \Pr(X > 320,000) \end{cases} &= \Pr\left(\frac{X-\mu}{\sigma} < \frac{180,000-\mu}{\sigma}\right) = \Phi\left(\frac{180,000-\mu}{\sigma}\right) \stackrel{(*)}{=} \Phi(-0.6744898) \\ &= \Pr\left(\frac{X-\mu}{\sigma} > \frac{320,000-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{320,000-\mu}{\sigma}\right) \stackrel{(*)}{=} 1 - \Phi(0.6744898) \\ \Rightarrow \begin{cases} \frac{180,000-\mu}{\sigma} = -0.6744898 \\ \frac{320,000-\mu}{\sigma} = 0.6744898 \end{cases} &\Rightarrow \mu = \frac{320,000 + 180,000}{2} = 250,000 \\ \Rightarrow \frac{180,000 - 250,000}{\sigma} = -0.6744898 &\Rightarrow \sigma = 103782.1 \end{aligned}$$

$$(★) : \Phi(-0.6744898) = 1 - \Phi(0.6744898) = 0.25$$

```
qnorm(0.25)
```

```
## [1] -0.6744898
```

(a)

$$Pr(X < 200,000) = Pr\left(\frac{X - \mu}{\sigma} < \frac{200,000 - 250,000}{103782.1}\right) \approx \Phi(-0.4817786) \approx 0.3149816$$

```
pnorm((200000-250000)/103782.1)
```

```
## [1] 0.3149816
```

(b)

$$\begin{aligned} Pr(280,000 < X < 320,000) &= Pr\left(\frac{280,000 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{320,000 - \mu}{\sigma}\right) \\ &= Pr\left(\frac{280,000 - \mu}{\sigma} < Z < \frac{320,000 - \mu}{\sigma}\right), \text{ where } Z \sim N(0, 1) \\ &= \Phi\left(\frac{320,000 - \mu}{\sigma}\right) - \Phi\left(\frac{280,000 - \mu}{\sigma}\right) \\ &\approx 0.75 - 0.613735 \\ &= 0.136265 \end{aligned}$$

```
0.75-pnorm((280000-250000)/103782.1)
```

```
## [1] 0.136265
```

Problem 29

Let X be the number of times the stock price goes up (i.e., the price becomes $u \times$ its original value) in the next 1000 periods.

Then, $X \sim \text{Binomial}(1000, p)$, and $(1000 - X)$ is the number of times the stock price goes down (i.e., the price becomes d times its original value) in the next 1000 periods.

When the initial stock price is s , after 1000 periods, the price becomes $s \times u^X \times d^{1000-X} = s \times d^{1000} \times \left(\frac{u}{d}\right)^X$. So the event of interest is :

$$(u^x d^{1000-x})s \geq 1.3s \iff (1.012^x \times 0.99^{1000-x}) \geq 1.3 \iff 1000 \log(0.99) + x \left(\log\left(\frac{1.012}{0.99}\right) \right) \geq \log(1.3) \iff x \geq 470.$$

Because $X \sim \text{Binomial}(1000, 0.52) \stackrel{d}{\approx} N(520, 249.6)$ by the [†]*DeMoivre – Laplace limit theorem* (see TBp.219),

$$Pr(X \geq 470) \approx Pr\left(\frac{X - 520}{\sqrt{249.6}} > \frac{470 - 0.5 - 520}{\sqrt{249.6}}\right) \approx Pr(Z > -3.196459) \approx 0.9993044$$

(Note that the "continuity correction" has been taken!)

[†]**FYI:** This result can also be obtained by applying the well-known Central Limit Theorem!

```
1-pnorm((470-0.5-520)/sqrt(249.6))
```

```
## [1] 0.9993044
```

Problem 30

Let X = the reading of the randomly chosen point.

Let $Y = \begin{cases} 0, & \text{if this point is in the white section} \\ 1, & \text{if this point is in the black section} \end{cases}$, then $Y \sim \text{Bernoulli}(\alpha) \iff p_Y(y) = \begin{cases} 1 - \alpha, & \text{if } y = 0 \\ \alpha, & \text{if } y = 1 \end{cases}$.

So we have $\begin{cases} (X | Y = 0) \sim N(4, 2^2) \\ (X | Y = 1) \sim N(6, 3^2) \end{cases}$ and we want to find α such that $Pr(Y = 0 | X = 5) = Pr(Y = 1 | X = 5) = \frac{1}{2}$.

$$\begin{aligned}
 \text{That is, } \frac{1}{2} &= Pr(\text{the chosen point having a reading of 5 is from the black section}) \\
 &= Pr(Y = 1 | X = 5) \\
 &= \frac{Pr(Y = 1, X = 5)}{Pr(X = 5)} \\
 &= \frac{Pr(Y = 1)Pr(X = 5 | Y = 1)}{Pr(Y = 1)Pr(X = 5 | Y = 1) + Pr(Y = 0)Pr(X = 5 | Y = 0)} \\
 &= \frac{\alpha \times \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{(5-4)^2}{2 \times 4}\right)}{\alpha \times \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{(5-4)^2}{2 \times 4}\right) + (1 - \alpha) \times \frac{1}{3\sqrt{2\pi}} \exp\left(-\frac{(5-6)^2}{2 \times 9}\right)} \\
 &\Rightarrow \dagger \alpha \approx 0.41677
 \end{aligned}$$

\dagger : From the above derivation, using $Pr(Y = 0 | X = 5) = \frac{1}{2}$ would also yield the same result.

(INTENTIONALLY LEFT BLANK)

Problem 31**(a)**

$$\begin{aligned}
E(|X - a|) &= \int_0^A |x - a| \frac{1}{A} dx = \frac{1}{A} \left(\int_0^a (a - x) dx + \int_a^A (x - a) dx \right) \\
&= \frac{1}{A} \left(\left(ax - \frac{x^2}{2} \right) \Big|_0^a + \left(\frac{x^2}{2} - ax \right) \Big|_a^A \right) = \frac{A}{2} - a + \frac{a^2}{A}
\end{aligned}$$

To find $\operatorname{argmin}_{a \in (0, A)} E(|X - a|)$, set $\left(\frac{d}{da} E(|X - a|) \right) \Big|_{a^*} = -1 + \frac{2a^*}{A} = 0$, then $a^* = \frac{A}{2}$.

Note that $\left(\frac{d^2}{da^2} E(|X - a|) \right) \Big|_{a^*} = \frac{2}{A} > 0$, so $a = \frac{A}{2}$ is actually the minimizer.

(b)

$$\begin{aligned}
E(|X - a|) &= \int_0^\infty |x - a| \lambda e^{-\lambda x} dx = \left(\int_0^a (a - x) \lambda e^{-\lambda x} dx + \int_a^\infty (x - a) \lambda e^{-\lambda x} dx \right) \\
&= \underbrace{\left(\int_0^a a \lambda e^{-\lambda x} dx \right)}_{(1)} - \underbrace{\left(\int_0^a \lambda x e^{-\lambda x} dx \right)}_{(2)} + \underbrace{\left(\int_a^\infty \lambda x e^{-\lambda x} dx \right)}_{(3)} - \underbrace{\left(\int_a^\infty a \lambda e^{-\lambda x} dx \right)}_{(4)} \\
&= \underbrace{\left(\left(-ae^{-\lambda x} \right) \Big|_0^a \right)}_{(1)} - \underbrace{\left(\left(xe^{-\lambda x} \right) \Big|_0^a + \int_0^a e^{-\lambda x} dx \right)}_{(2)} + \underbrace{\left(\left(-xe^{-\lambda x} \right) \Big|_a^\infty + \int_a^\infty e^{-\lambda x} dx \right)}_{(3)} + \underbrace{\left(\left(ae^{-\lambda x} \right) \Big|_a^\infty \right)}_{(4)} \\
&= \left(\cancel{-ae^{-\lambda a}} + a \right) - \left(\cancel{ae^{-\lambda a}} - \frac{1}{\lambda} e^{-\lambda a} + \frac{1}{\lambda} \right) + \left(\cancel{ae^{-\lambda a}} + \frac{1}{\lambda} e^{-\lambda a} \right) + \left(\cancel{ae^{-\lambda a}} \right) \\
&= a + \frac{2}{\lambda} e^{-\lambda a} - \frac{1}{\lambda}
\end{aligned}$$

$$\left\{ \begin{array}{l}
(1) : \text{Direct integration} \\
(2) : \text{Do integration by parts with } \left(\begin{array}{ll} u = x & du = dx \\ v = -e^{-\lambda x} & dv = \lambda e^{-\lambda x} dx \end{array} \right) \text{ over } \{x \in \mathcal{R} \mid 0 < x < a\} \\
(3) : \text{Do integration by parts with } \left(\begin{array}{ll} u = x & du = dx \\ v = -e^{-\lambda x} & dv = \lambda e^{-\lambda x} dx \end{array} \right) \text{ over } \{x \in \mathcal{R} \mid a < x\} \\
(4) : \text{Direct integration}
\end{array} \right.$$

To find $\operatorname{argmin}_{a \in (0, A)} E(|X - a|)$, set $\left(\frac{d}{da} E(|X - a|) \right) \Big|_{a^*} = 1 - 2e^{-\lambda a^*} = 0$, then $a^* = \frac{\log(2)}{\lambda}$.

Note that $\left(\frac{d^2}{da^2} E(|X - a|) \right) \Big|_{a^*} = 2\lambda e^{-\lambda a^*} > 0$, so $a = \frac{\log(2)}{\lambda}$ is actually the minimizer.

Problem 32

Because $X \sim \text{exponential}(\lambda = \frac{1}{1.5})$, the cdf of X is $F_X(x) = 1 - e^{-\frac{2}{3}x}$ for $x \geq 0$.

(a)

The desired quantity is $\Pr(X > 2) = 1 - F_X(2) = e^{-4/3}$.

(b)

The desired quantity is $\Pr(X > 2 \mid X > 1) = \frac{\Pr(X > 2)}{\Pr(X > 1)} = \frac{e^{-4/3}}{e^{-2/3}} = e^{-2/3}$, which equals $\Pr(X > 1)$.

This shows the *memoryless property* of an exponential distribution.

Theoretical Exercise 13

(a)

Let $X \sim \text{Uniform}(a, b)$, then the cdf of X is $F_X(x) = \begin{cases} \int_a^x \frac{1}{b-a} du & , \text{ if } a < x < b \\ 0 & , \text{ if } x \leq a \end{cases}$.

Let $m_{(a)}$ be the median of X , so we have :

$$F_X(m_{(a)}) = \frac{1}{2} = \int_a^{m_{(a)}} \frac{1}{b-a} du = \frac{u}{b-a} \Big|_{u=a}^{u=m_{(a)}} = \frac{m_{(a)} - a}{b-a} \Rightarrow m_{(a)} = \frac{a+b}{2}$$

(b)

Let $m_{(b)}$ be the median of X , where $X \sim N(\mu, \sigma^2)$.

Then the pdf of X , say $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$, is symmetric about μ ,

i.e., $f_X(\mu + \delta) = f_X(\mu - \delta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\delta^2}{2\sigma^2}\right\}$, $\forall \delta \in \mathcal{R} \Rightarrow m_{(b)} = \mu$ (*)

$$\textcolor{red}{(*)} : \int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\mu} f_X(x) dx + \int_{\mu}^{\infty} f_X(x) dx = 2 \int_{-\infty}^{\mu} f_X(x) dx = 1$$

(c)

Let $X \sim \text{Exp}(\lambda)$, then the cdf of X is $F_X(x) = \begin{cases} 1 - e^{-\lambda x} & , \text{ if } 0 < x \\ 0 & , \text{ if } x \leq 0 \end{cases}$.

Let $m_{(c)}$ be the median of X , then $\frac{1}{2} = 1 - e^{-\lambda m_{(c)}}$.

$$\Rightarrow \log\left(\frac{1}{2}\right) = -\lambda m_{(c)} \Rightarrow m_{(c)} = \frac{\log(2)}{\lambda}$$

Theoretical Exercise 19

The pdf of X is $f_X(x) = \lambda e^{-\lambda x}$, $x \geq 0$, so

$$\begin{aligned} E[X^k] &= \int_0^\infty x^k f_X(x) dx = \int_0^\infty \lambda x^k e^{-\lambda x} dx \\ &= \frac{\Gamma(k+1)}{\lambda^k} \underbrace{\int_0^\infty \frac{\lambda^{k+1}}{\Gamma(k+1)} x^{(k+1)-1} e^{-\lambda x} dx}_{(\star)} \\ &= \frac{\Gamma(k+1)}{\lambda^k} = \frac{k!}{\lambda^k} \end{aligned}$$

(Note that (\star) equals 1, since its integrand is the pdf of a $\Gamma(k+1, \lambda)$ distribution.)

(INTENTIONALLY LEFT BLANK)

Theoretical Exercise 31

We can find the pdf f_Y of Y by $\overbrace{\text{applying the theorem in LNp.6-10}}^{(*)}$.

Let $f_X(x)$ be the pdf of X , where $X \sim N(\mu, \sigma^2)$.

Let $Y = g(X) \equiv e^X$, so the range of Y is $R_Y = \{y \mid y > 0\}$.

We then have $g^{-1}(Y) = \log(X)$ and $\left| \frac{dg^{-1}(y)}{dy} \right| = \frac{1}{y}$.

$(*)$: (Note that $\frac{d}{dx}g(x) = e^x > 0 \forall x \in R_X = \mathcal{R}$, so g is differentiable and strictly monotone.)

So the pdf of y is:

$$f_Y(y) = \begin{cases} f_X(g^{-1}(x)) \left| \frac{dg^{-1}(y)}{dy} \right| = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(\log(y) - \mu)^2}{2\sigma^2} \right\} & , y > 0 \\ 0 & , y \leq 0 \end{cases}.$$

An alternative approach is to first derive the cdf F_Y of Y , and then obtain the pdf f_Y from the cdf F_Y as follows.

$$(**) \left\{ \begin{aligned} F_Y(y) &= Pr\{Y \leq y\} = Pr\{e^X \leq y\} \\ &= \begin{cases} Pr\{X < \log(y)\} & , \text{ if } y > 0 \\ 0 & , \text{ if } y \leq 0 \end{cases} \\ &= \begin{cases} \int_{-\infty}^{\log(y)} f_X(x) dx & , \text{ if } y > 0 \\ 0 & , \text{ if } y \leq 0 \end{cases} \end{aligned} \right.$$

Take derivative on $(**)$ with respect to y over $\{y \mid y > 0\}$, we obtain:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{y} f_X(\log(y)) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(\log(y) - \mu)^2}{2\sigma^2} \right\}, \text{ if } y > 0.$$

(Note that we have applied the † Leibniz integral rule.)

Also, it is obvious that $f_Y(y) = 0$, if $y < 0$.

Thus, the pdf of Y , $f_Y(y)$, is:

$$f_Y(y) = \begin{cases} \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(\log(y) - \mu)^2}{2\sigma^2} \right\} & , \text{ if } y > 0 \\ 0 & , \text{ if } y \leq 0 \end{cases}$$

This verifies our answer above!

† : You may browse this page for details about Leibniz integral rule.

R

The following is a brief introduction to R language (about pdf, cdf, quantiles, generating random samples).

Some problems above require finding the quantiles of a normal distribution, but looking them up on tables is a very outdated practice. Using R is a more efficient method, below are some basic functions of R:

Let $X \sim N(0, 1)$, which is the same as default.

1. $\text{pnorm}(3) = Pr(X \leq 3) = \Phi(3)$, is the cdf.
2. $\text{qnorm}(0.2) = \Phi^{-1}(0.2) = z_{0.2}$.
3. $\text{dnorm}(0.4) = f_X(0.4)$ is the pdf of X at 0.4.
4. $\text{rnorm}(5)$ generates a sample iid from $N(0,1)$ of size “5”.

For example:

- To have a $N(\mu, \sigma^2)$ distribution, use $\text{norm}(\mu, \sigma)$, for example, $\text{norm}(3,2)$ is $N(3,4)$.
- To generate a sample iid from $N(3,4)$ of size 100, use $\text{rnorm}(100,3,2)$.
- To find $z_{0.05}$, use $\text{qnorm}(0.05)$.

The above (p,q,d,r) functions can also be used for other distributions, for example:

$$\left\{ \begin{array}{l} \text{(What we want)} \leftrightarrow \text{(Code)} \\ \text{Negative Binomial} \leftrightarrow \text{nbinom} \\ \text{Hypergeometric} \leftrightarrow \text{hyper} \\ \text{Exponential} \leftrightarrow \text{exp} \\ \text{Geometric} \leftrightarrow \text{geom} \\ \text{Uniform} \leftrightarrow \text{unif} \\ \text{Binomial} \leftrightarrow \text{binom} \\ \text{Normal} \leftrightarrow \text{norm} \\ \text{Gamma} \leftrightarrow \text{gamma} \\ \text{Poisson} \leftrightarrow \text{pois} \\ \text{Weibull} \leftrightarrow \text{weibull} \\ \text{Cauchy} \leftrightarrow \text{cauchy} \\ \text{Beta} \leftrightarrow \text{beta} \\ \chi^2 \leftrightarrow \text{chisq} \\ \mathcal{T} \leftrightarrow t \\ \mathcal{F} \leftrightarrow f \end{array} \right\} \Rightarrow \dagger \text{Use } (p, q, r, d) + \text{RHS, like rt, qbeta, dchisq, or pf} \dots$$

\dagger : You must enter the required parameters, otherwise R will use the default parameters.

Resources with hyperlinks:

A resource for those who have not used R, including instructions for downloading R!

An entry-level learning resource!