STAT 509 Homework 3 Solution

Chong Ma

February 13, 2017

Problem 1

Note that Y = amount of sapce in (ft^3) occupied in this container and the pdf of Y is

$$f_Y(y) = \begin{cases} 90y^8(1-y), & 0 < y < 1 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Since

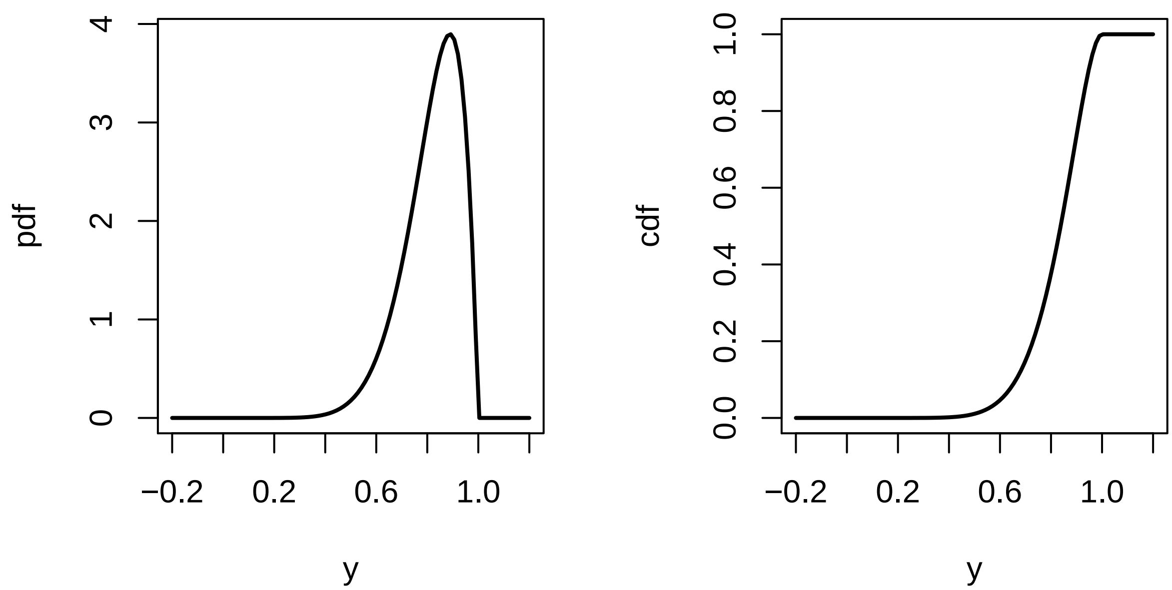
$$\int_0^y 90t^8(1-t)dt = 90\Big(\frac{t^9}{9} - \frac{t^{10}}{10}\Big)\Big|_0^y = 10y^9(1-0.9y)$$

then the cdf of Y is obtained by

$$F_Y(y) = \int_{-\infty}^y f_Y(t)dt = \begin{cases} 0, & y \leq 0 \\ 10y^9(1-0.9y), & 0 < y < 1 \\ 1, & y \geq 1 \end{cases}$$

(a) Graph the pdf of Y

```
pdf<-function(y) ifelse(y>0 & y<1,90*y^8*(1-y),0)
cdf<-function(y) ifelse(y<=0, 0, ifelse(y<1, 10*y^9*(1-0.9*y),1))
par(mfrow=c(1,2))
curve(pdf,-0.2,1.2,xlab="y",ylab="pdf",lwd=2)
curve(cdf,-0.2,1.2,xlab="y",ylab="cdf",lwd=2)
```



(b) The probability that the contents of a randomly selected 1-ft^3 packing container will occupy less than 1/2 of the volume of the container is calculated by

$$P(Y < 0.5) = F_Y(0.5) = 10(0.5)^9(1 - 0.9(0.5)) = 0.0107422$$

(c) The expected value and variance of Y exist in closed form, which are

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y f_Y(y)dy = \int_0^1 y(90y^8(1-y))dy \\ &= 90\Big(\frac{1}{10}y^10 - \frac{1}{11}y^11\Big)\Big|_0^1 = \frac{9}{11} \end{aligned}$$

Note that

$$\begin{aligned} E(Y^2) &= \int_{-\infty}^{\infty} y^2 f_Y(y)dy = \int_0^1 y^2(90y^8(1-y))dy \\ &= 90\Big(\frac{1}{11}y^11 - \frac{1}{12}y^12\Big)\Big|_0^1 = \frac{90}{132} \end{aligned}$$

thus, $Var(Y) = E(Y^2) - [E(Y)]^2 = \frac{90}{132} - (\frac{9}{11})^2 = 0.0123967$

(d) To obtain the median, that is $\phi_{0.5}$ of Y, it solves the following eqaution

$$F_Y(\phi_{0.5}) = 10{\phi_{0.5}}^9(1 - 0.9\phi_{0.5}) = 0.5$$

Since the equation involves a 10 degree polynomial with only one real root in (0,1), we need solve it numerically. By using **uniroot** R fucntion, we get the median of Y is $\phi_{0.5} = 0.8377$.

```r
## get the objective function
obj<-function(y) cdf(y) - 0.5
## fint the root of F(y)-0.5 in the interval (0,1)
uniroot(obj,c(0,1))
```

```
## $root
## [1] 0.8377382
##
## $f.root
## [1] 3.418859e-06
##
## $iter
## [1] 7
##
## $init.it
## [1] NA
##
## $estim.prec
## [1] 6.103516e-05
```

Problem 2

Note that $Y \sim Uniform(a, b)$, the pdf of Y is

$$f_Y(y) = \begin{cases} \frac{1}{b-a}, & a < y < b \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

(a) The cdf of Y is

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \begin{cases} 0, & y \leq a \\ \frac{y-a}{b-a}, & a < y < b \\ 1, & y \geq b \end{cases}$$

(b) Let us find the general formula for expected value and variance of $Y \sim Uniform(a, b)$. The expected value of Y is

$$E(Y) = \int_{-\infty}^{\infty} \frac{1}{b-a} y dy = \int_{a}^{b} \frac{1}{b-a} y dy = \frac{1}{b-a}(\frac{1}{2}y^2)|_a^b = \frac{a+b}{2}$$

Also note that

$$E(Y^2) = \int_{-\infty}^{\infty} \frac{1}{b-a} y^2 dy = \int_{a}^{b} \frac{1}{b-a} y dy = \frac{1}{b-a}(\frac{1}{3}y^3)|_a^b = \frac{b^3 - a^3}{3(b-a)}$$

therefore, the variance of Y is

$$Var(Y) = E(Y^2) - [E(Y)]^2 = \frac{b^3 - a^3}{3(b-a)} - \Big(\frac{a+b}{2}\Big)^2 = \frac{(b-a)^2}{12}$$

When $a = 0.01$ and $b = 0.05$, using the derived formula,

(i) The mean and variance of $Y \sim Uniform(0.01, 0.05)$ is

$$\begin{aligned} E(Y) &= \frac{a+b}{2} = \frac{0.01 + 0.05}{2} = 0.03 \\ Var(Y) &= \frac{(b-a)^2}{12} = \frac{(0.05 - 0.01)^2}{12} = 10^{-4} \end{aligned}$$

(ii) The probability that a particle has diameter larger than 0.045 cm is

$$P(Y > 0.045) = 1 - P(Y \leq 0.045) = 1 - F_Y(0.045) = 1 - \frac{0.045 - 0.05}{0.05 - 0.01} = 0.125$$

Problem 3

Note that Y = the time until a major overhaul is needed and the pdf of Y is given by

$$f_Y(y) = ce^{-y/2} I(y > 0)$$

(a) Because $f_Y(y)$ is a pdf, then $\int_{-\infty}^{\infty} f_Y(y)dy = 1$. Hence,

$$\int_{-\infty}^{\infty} f_Y(y)dy = \int_{-\infty}^{\infty} ce^{-y/2}dy = 2c = 1$$

That is, $c = 1/2$.

(b) From part (a), we know that $Y \sim exponential(\lambda = 1/2)$. Therefore, the mean and variance of Y are

$$E(Y) = \frac{1}{\lambda} = 2$$
$$Var(Y) = \frac{1}{\lambda^2} = 4$$

(c) The probability that the time to a major overhaul is longer than 3 years is

$$P(Y > 3) = 1 - P(Y \leq 3) = 1 - F_Y(3) = 1 - (1 - e^{-\frac{1}{2}(3)}) = 0.223$$

Since the probability is about 0.223, indiating that 3 years is the 22th percentile of the time to a major overhaul, it should not be considered unusual.

(d) To get the 95th percentile of this distribution is to solve the following equation

$$F_Y(\phi_{0.95}) = 1 - e^{-\frac{1}{2}\phi_{0.95}} = 0.95$$

Therefore, $\phi_{0.95} = -2log(0.05) = 5.9915$

(e) For a general exponential distribution $Y \sim exponential(\lambda)$, the **moment generation function(mgf)** of Y is

$$\begin{aligned} E(e^{tY}) &= \int_0^{\infty} e^{ty}\lambda e^{-\lambda y}dy = \frac{\lambda}{\lambda - t}\int_0^{\infty}(\lambda - t)e^{-(\lambda - t)y}dy \\ &= \frac{\lambda}{\lambda - t} = \left(1 - \frac{t}{\lambda}\right)^{-1} \quad \text{for } t < \lambda \end{aligned}$$

since $\int_0^{\infty}(\lambda - t)e^{-(\lambda - t)y}dy = 1$ (the integrand is essentially a pdf that requires $\lambda - t > 0$, that is, $t < \lambda$). Therefore, when $\lambda = 1/2$, the **mgf** of Y turns out

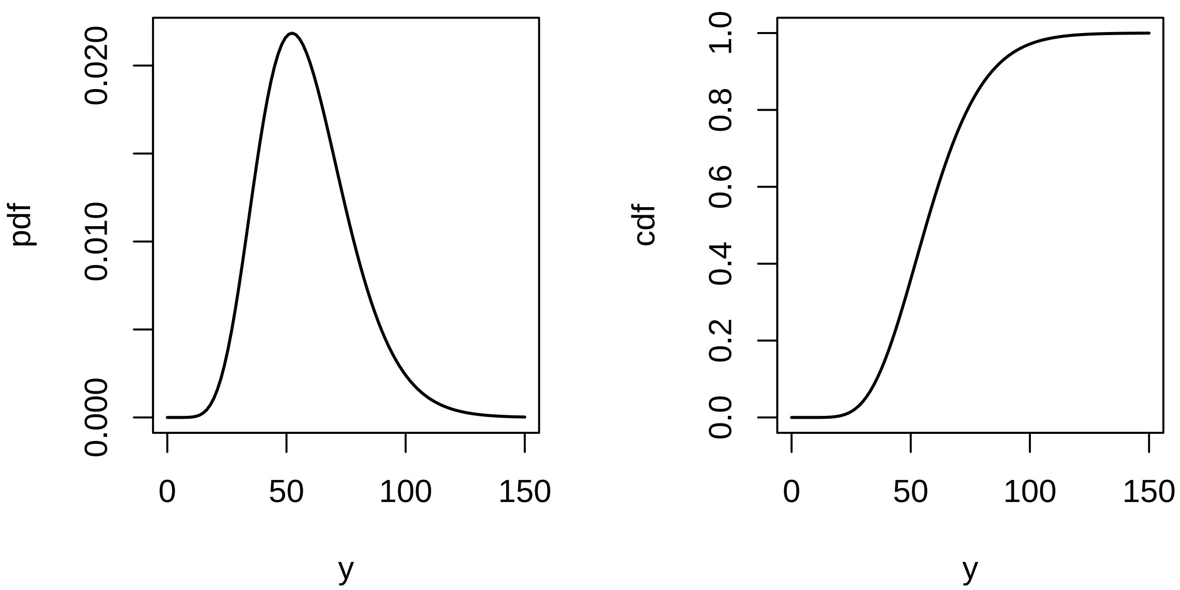
$$E(e^{tY}) = \frac{1}{1 - 2t}$$

Problem 4

Note that Y = the alkalinity concentration in ground water and $Y \sim Gamma(\alpha = 9.37, \lambda = 0.16)$.

(a) The **probability density function(pdf)** and **cumulative distribution function(cdf)** of Y are given by

```
par(mfrow=c(1,2))
curve(dgamma(x,9.37,0.16),0,150,lwd=1.5,xlab="y",ylab="pdf")
curve(pgamma(x,9.37,0.16),0,150,lwd=1.5,xlab="y",ylab="cdf")
```



(b) Recall the mean of Y is $E(Y) = \frac{\alpha}{\lambda^2}$, thus

$$E(Y) = \frac{\alpha}{\lambda^2} = \frac{9.37}{0.16^2} = 366.0156$$

And the median of Y is $qgamma(0.5, 9.37, 0.16) = 56.49$

(c) The proportion of ground water samples near this landfill will have alkaliity concentrations less than 80 mg/L is

$$P(Y < 80) = \int_0^{80} \frac{0.16^{9.37}}{\Gamma(9.37)} y^{9.37-1} e^{-0.16y} dy = pgamma(80, 9.37, 0.16) = 0.87$$

(d) To find the 99th percentile of the distribution is to solve the following eqaution

$$F_Y(\phi_{0.99}) = \int_0^{\phi_{0.99}} \frac{0.16^{9.37}}{\Gamma(9.37)} y^{9.37-1} e^{-0.16y} dy = 0.99$$

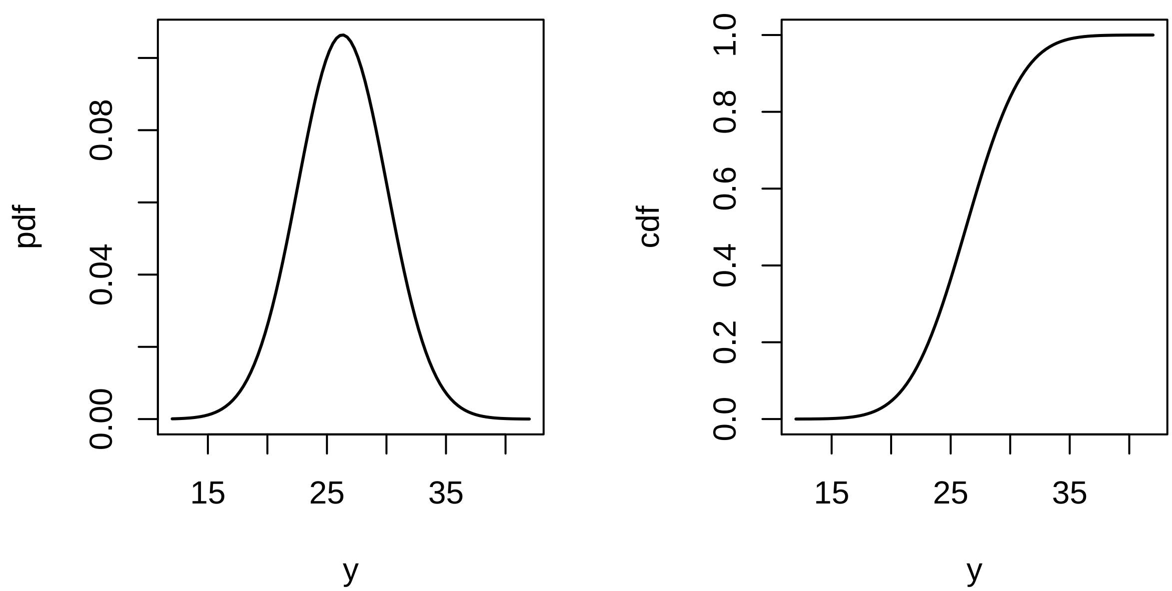
Thus, $\phi_{0.99} = qgamma(0.99, 9.37, 0.16) = 111.97$. It means that only one percent of ground water samples near this landfill will have alkalinity concentration more than 111.97(mg/L).

Problem 5

Note that Y = the age at which the subject gave birth to her first child and $Y \sim Normal(\mu = 26.3, \sigma = 3.75)$.

(a) Graph the pdf and cdf of Y

```
par(mfrow=c(1,2))
curve(dnorm(x,26.3,3.75),12,42,xlab="y",ylab="pdf",lwd=1.5)
curve(pnorm(x,26.3,3.75),12,42,xlab="y",ylab="cdf",lwd=1.5)
```



(b) The probability that a mother gives birth to her first child after she is 35 years old is

$$P(Y > 35) = 1 - P(Y \leq 35) = 1 - \int_{-\infty}^{35} \frac{1}{\sqrt{2\pi(3.75)}} e^{-\frac{1}{2}(\frac{y-26.3}{3.75})^2} dy = 1 - \text{pnorm(35,26.3,3.75)} = 0.01$$

(c) The proportion of first-time mothers give birth between 20 and 30 years is

$$P(20 < Y < 30) = \text{pnorm(30,26.3,3.75)} - \text{pnorm(20,26.3,3.75)} = 0.79$$

(d) To get the 10th percentile of the age at which first-time mothers give birth is to solve the following function,

$$F(\phi_{0.1}) = \int_{-\infty}^{\phi_{0.1}} \frac{1}{\sqrt{2\pi(3.75)}} e^{-\frac{1}{2}(\frac{y-26.3}{3.75})^2} dy = 0.1$$

that is, $\phi_{0.1} = qnorm(0.1, 26.3, 3.75) = 21.49$