STAT2001 & STAT6039 Final Exam June 2016 Solutions

Solution to Problem 1

(a) X has pdf $f(x) = 2x, 0 < x < 1$. Also, $y = x/(x+1)$ is a strictly increasing function with inverse $x = y(1-y)^{-1}$. So, by the transformation rule,

$$f(y) = f(x)\left|\frac{dx}{dy}\right| = 2 \times y(1-y)^{-1} \times \left| y(-1)(1-y)^{-2}(-1) + 1 \times (1-y)^{-1} \right|$$

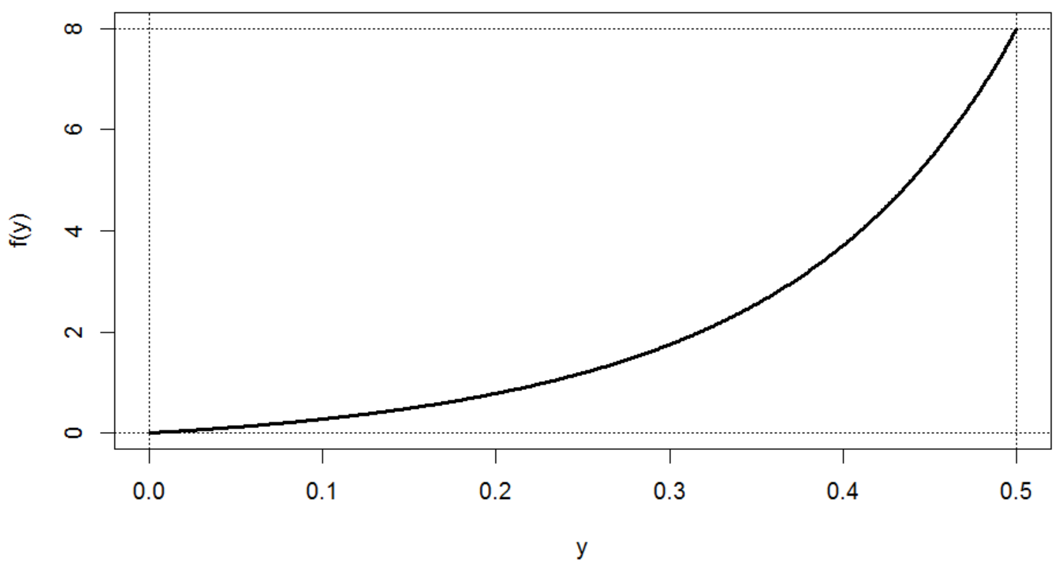
$$= \frac{2y}{1-y}\left(\frac{y}{(1-y)^2} + \frac{1}{1-y}\right).$$

Thus, $\boxed{f(y) = \frac{2y}{(1-y)^3}, 0 < y < \frac{1}{2}}$, as illustrated in the sketch below.

Also, $EY = E\left(\frac{X}{X+1}\right) = \int_0^1 \frac{x}{x+1} \times 2x dx = 2\int_1^2 \frac{(t-1)^2}{t} dt$ (where $t = x+1$)

$$= 2\int_1^2 \left(t - 2 + \frac{1}{t}\right) dt = 2\left[\frac{t^2}{2} - 2t + \log t \Big|_{t=1}^{2}\right] = 2\left(\frac{2^2}{2} - 2 \times 2 + \log 2 - \frac{1^2}{2} + 2 \times 1 - \log 1\right)$$

$$= 2\left(2 - 4 + \log 2 - \frac{1}{2} + 2\right) = 2\log 2 - 1 = \boxed{0.3863}.$$



(b) R has cdf $\quad F(r) = P(R \le r) = P(Z/U \le r) = P(Z \le rU)$

$$= \int_{u=0}^{1} \left(\int_{z=-\infty}^{ru} \phi(z)dz \right) 1du = \int_{u=0}^{1} \Phi(ru)du\,. \qquad (1)$$

So R has pdf $\quad f(r) = F'(r) = \int_{u=0}^{1} \phi(ru)udu = \int_{0}^{1} \frac{1}{\sqrt{2\tau}} e^{-\frac{1}{2}(ru)^2} udu\,.$

We now substitute $t = \frac{1}{2}r^2u^2$, so that $\frac{dt}{du} = r^2u$ and $udu = \frac{dt}{r^2}$,

and thereby obtain $\quad f(r) = \frac{1}{r^2\sqrt{2\pi}} \int_{0}^{\frac{1}{2}r^2} e^{-t}dt\,.$

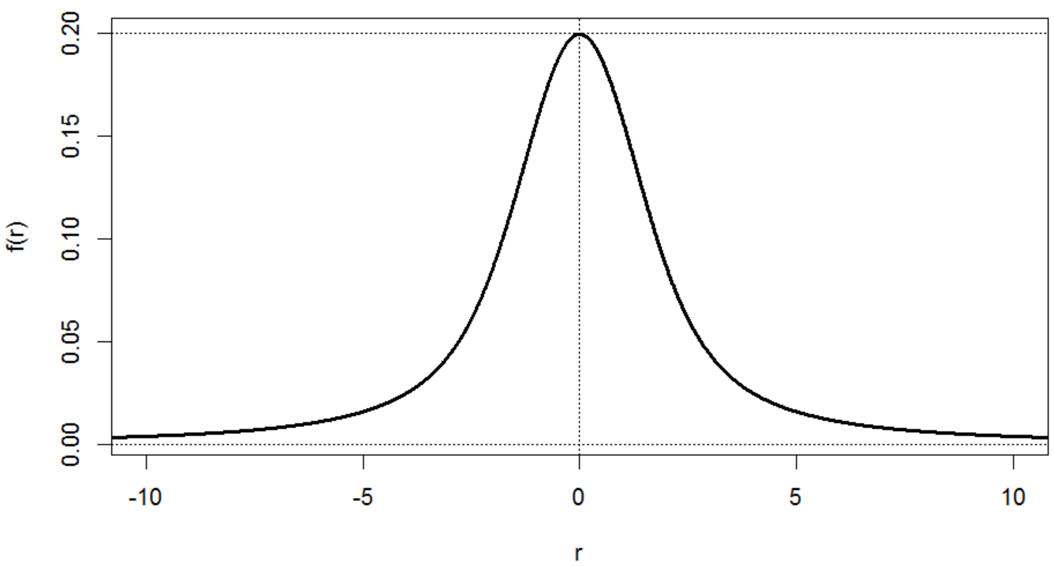
Thus $\boxed{f(r) = \frac{1}{r^2\sqrt{2\pi}} \left(1 - e^{-\frac{1}{2}r^2}\right), \quad r \in \Re \ \ (r \ne 0)}$, as shown in the sketch below.

We see that $f(r)$ is symmetric around zero, with a maximum of approximately

$$\frac{1}{0.0001^2\sqrt{2\pi}} \left(1 - e^{-\frac{1}{2}0.0001^2}\right) = 0.1994711 \approx 0.2\,.$$

Note: More accurately, we can apply L'Hospital's rule to get

$$\lim_{r\to 0} \left\{ \frac{1}{r^2\sqrt{2\pi}} \left(1 - e^{-\frac{1}{2}r^2}\right) \right\} = \lim_{r\to 0} \left\{ \frac{1}{2\cancel{r}\sqrt{2\pi}} \left(0 - e^{-\frac{1}{2}r^2}(-\cancel{r})\right) \right\} = \frac{1}{\sqrt{8\pi}} = 0.1994711.$$



To find $P(R>8)$, we consider any $r > 0$ (e.g. 8) and the region under the line $z = ru$ in the *u*-*z* plane within the infinite rectangle $(0,1)\times(-\infty,\infty)$. We then write (1) as

$$F(r)=\int_{u=0}^{1}\left(\int_{z=-\infty}^{ru}\phi(z)dz\right)1du=\frac{1}{2}+\int_{z=0}^{r}\left(\int_{u=z/r}^{1}1du\right)\phi(z)dz$$

$$=\frac{1}{2}+\int_{z=0}^{r}\left(1-\frac{z}{r}\right)\phi(z)dz=\frac{1}{2}+\int_{0}^{r}\phi(z)dz-\frac{1}{r}\int_{0}^{r}z\phi(z)dz$$

$$=\frac{1}{2}+\left(\Phi(r)-\frac{1}{2}\right)-\frac{1}{r}I,\quad\text{where } I=\int_{0}^{r}z\phi(z)dz=\int_{0}^{r}z\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz.$$

We now substitute $t=\frac{1}{2}z^2$, so that $\frac{dt}{dz}=z$ and $zdz=dt$.

Thereby we obtain $I=\frac{1}{\sqrt{2\pi}}\int_{0}^{\frac{1}{2}r^2}e^{-t}dt=\frac{1}{\sqrt{2\pi}}\left(1-e^{-\frac{1}{2}r^2}\right)$,

and hence $F(r)=\Phi(r)-\frac{1}{r}\left\{\frac{1}{\sqrt{2\pi}}\left(1-e^{-\frac{1}{2}r^2}\right)\right\}$ $(r > 0)$. (2)

It follows that

$$P(R>8)=1-F_R(8)=1-\Phi(r)+\frac{1}{r}\left\{\frac{1}{\sqrt{2\pi}}\left(1-e^{-\frac{1}{2}r^2}\right)\right\}\quad\text{where } r=8$$

$$\approx\frac{1}{8}\left\{\frac{1}{\sqrt{2\pi}}\left(1-e^{-\frac{1}{2}8^2}\right)\right\}\quad(\text{since } \Phi(8)=P(Z\le 8)\approx 1)$$

$$=\boxed{0.04987}.$$

Note: Equation (2) provides an alternative way to get the pdf of *R*, namely as

$$f(r)=F'(r)=\phi(r)-\frac{1}{\cancel{r}}\left\{\frac{1}{\sqrt{2\pi}}\left(0-e^{-\frac{1}{2}r^2}(-\cancel{r})\right)\right\}-\left(\frac{-1}{r^2}\right)\left\{\frac{1}{\sqrt{2\pi}}\left(1-e^{-\frac{1}{2}r^2}\right)\right\}$$

$$=\cancel{\phi(r)}-\cancel{\phi(r)}+\frac{1}{r^2\sqrt{2\pi}}\left(1-e^{-\frac{1}{2}r^2}\right),\ r>0\text{ (as before).}$$

By symmetry, this last formula must also give the pdf of *R* when $r < 0$.

R Code for Problem 1 (not required)

```
# (a)
yvec=seq(0,0.5,0.001); fyvec=2*yvec/(1-yvec)^3
X11(w=8,h=5); plot(yvec,fyvec,type="l",lwd=3,xlab="y",ylab="f(y)")
abline(h=c(0,8), lty=3); abline(v=c(0,0.5),lty=3)
2*log(2)-1 # 0.3862944
# Check mean via Monte Carlo
set.seed(331); xv=rbeta(100000,2,1); yv=xv/(xv+1); mean(yv) # 0.3860128 OK

# (b)
rvec=seq(-12.001,11.999,0.01)
frvec= (1/ (rvec^2* sqrt(2*pi) ) ) * (1-exp(-0.5*rvec^2))
rval=0.0001; (1/ (rval^2* sqrt(2*pi) ) ) * (1-exp(-0.5*rval^2)) # 0.1994711
1/sqrt(8*pi) # 0.1994711 (exact mode)
X11(w=8,h=5); plot(rvec,frvec,type="l",lwd=3,xlab="r",ylab="f(r)", xlim=c(-10,10))
abline(h=c(0,0.2), lty=3); abline(v=0,lty=3)
rval=8; (1/rval)*(1/(sqrt(2*pi)))*(1-exp(-0.5*rval^2)) # 0.04986779
# Check probability via Monte Carlo
set.seed(768); uv=runif(100000); zv=rnorm(100000); rv=zv/uv
length(rv[rv>8])/100000 # 0.0491 OK
```

Solution to Problem 2

(a) Let N ="Number of rolls until the first swipe". Then, by a first-step analysis:

$$EN = P(1)E(N|1) + P(2)E(N|2) + P(3)E(N|3) + P(0)E(N|0)$$

where 0 = "4, 5 or 6"

$$= \frac{1}{6}E(N|1) + \frac{1}{6}E(N|2) + \frac{1}{6}E(N|3) + \frac{3}{6}\{EN+1\}$$

$$= \frac{3}{6}E(N|1) + \frac{3}{6}\{EN+1\} \qquad \text{since } E(N|1) = E(N|2) = E(N|3).$$

Also,

$$E(N|1) = P(11|1)E(N|1,11) + P(12|1)E(N|1,12)$$
$$+P(13|1)E(N|1,13) + P(10|1)E(N|1,10)$$
$$= \frac{1}{6}E(N|11) + \frac{1}{6}E(N|12) + \frac{1}{6}E(N|13) + \frac{3}{6}\{EN+2\}$$
$$= \frac{1}{6}\{E(N|1)+1\} + \frac{2}{6}E(N|12) + \frac{3}{6}\{EN+2\} \text{ since } \quad E(N|13) = E(N|12).$$

Furthermore,

$$E(N|12) = P(121|12)E(N|12,121) + P(122|12)E(N|12,122)$$
$$+P(123|12)E(N|12,123) + P(120|12)E(N|12,120)$$
$$= \frac{1}{6}\{E(N|12)+1\} + \frac{1}{6}\{E(N|1)+2\} + \frac{1}{6}\times 3 + \frac{3}{6}\{EN+3\}.$$

Writing $a = EN$, and $b = E(N|1)$, $c = E(N|12)$, these equations may be written as:

$$6a = 3b + 3a + 3 \quad \Rightarrow \quad b = a - 1 \qquad (1)$$
$$6b = b + 1 + 2c + 3a + 6 \quad \Rightarrow \quad 5b = 2c + 3a + 7 \qquad (2)$$
$$6c = c + 1 + b + 2 + 3 + 3a + 9 \quad \Rightarrow \quad 5c = b + 3a + 15. \qquad (3)$$

Then: $\quad$ (1) $\rightarrow$ (2) $\Rightarrow 5(a-1) = 2c + 3a + 7 \quad \Rightarrow c = a - 6$ $\qquad$ (4)

(1) & (4) $\rightarrow$(3) $\Rightarrow$ $5(a-6) = a - 1 + 3a + 15 \Rightarrow a = \boxed{44}$.

(b) Let A ="Number of rolls to first swipe is even". Then, by a first-step analysis:

$$P(A) = P(1)P(A|1) + P(2)P(A|2) + P(3)P(A|3) + P(0)P(A|0)$$

$$= \frac{1}{6}P(A|1) + \frac{1}{6}P(A|2) + \frac{1}{6}P(A|3) + \frac{3}{6}\{1 - P(A)\}$$

$$= \frac{3}{6}P(A|1) + \frac{3}{6}\{1 - P(A)\} \qquad \text{since } P(A|1) = P(A|2) = P(A|3).$$

Also,

$$P(A|1) = P(11|1)P(A|1,11) + P(12|1)P(A|1,12)$$

$$+P(13|1)P(A|1,13) + P(10|1)P(A|1,10)$$

$$= \frac{1}{6}\{1 - P(A|1)\} + \frac{1}{6}P(A|12) + \frac{1}{6}P(A|13) + \frac{3}{6}P(A)$$

$$= \frac{1}{6}\{1 - P(A|1)\} + \frac{2}{6}P(A|12) + \frac{3}{6}P(A) \text{ since } P(A|13) = P(A|12).$$

Furthermore,

$$P(A|12) = P(121|12)P(A|12,121) + P(122|12)P(A|12,122)$$

$$+P(123|12)P(A|12,123) + P(120|12)P(A|12,120)$$

$$= \frac{1}{6}\{1 - P(A|21)\} + \frac{1}{6}P(A|2) + \frac{1}{6}\times 0 + \frac{3}{6}\{1 - P(A)\}$$

$$= \frac{1}{6}\{1 - P(A|12)\} + \frac{1}{6}P(A|1) + \frac{3}{6}\{1 - P(A)\} \text{ since } P(A|21) = P(A|12).$$

Writing $a = P(A)$, and $b = P(A|1)$, $c = P(A|12)$, these equations may be written as:

$$6a = 3b + 3 - 3a \quad \Rightarrow \quad b = 3a - 1 \qquad (1)$$

$$6b = 1 - b + 2c + 3a \quad \Rightarrow \quad 7b = 1 + 2c + 3a \qquad (2)$$

$$6c = 1 - b + 2c + 3a \quad \Rightarrow \quad 4c = 1 - b + 3a. \qquad (3)$$

Then: $(1) \to (2) \Rightarrow 21a - 7 = 1 + 2c + 3a \quad \Rightarrow c = 9a - 4$ (4)

$(1)\ \&\ (4) \to (3) \Rightarrow 36 - 16 = 1 - 3a + 1 + 3a \Rightarrow a = \boxed{1/2}$.

(c) Let A_i = "Swipe on rolls i, i + 1 and i + 2" and $Y_i = I(A_i)$.

Then, on n rolls, the number of swipes is $Y = Y_1 + ... + Y_{n-2}$, with expectation

$$EY = EY_1 + ... + EY_{n-2} = (n-2)EY_1 = (n-2)P(A_1) = (n-2)\frac{3}{6}\times\frac{2}{6}\times\frac{1}{6} = \frac{n-2}{36}.$$

Setting EY to 1 yields $n = \boxed{38}$.

(d) The probability of at least one swipe on six rolls is

$$p = P(A_1 \cup \ldots \cup A_4) = \sum_{i=1,\ldots,4} P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) - P(A_1 \ldots A_4) .$$

Now: $P(A_1) = \dfrac{3!}{6^3} = P(A_2) = P(A_3) = P(A_4) = \dfrac{216}{d}$ where $d = 6^5 = 7776$

$$P(A_1 A_2) = \frac{3!}{6^3} \times \frac{1}{6} = P(A_{23}) = P(A_{34}) = \frac{36}{d}$$

$$P(A_1 A_3) = \frac{3!}{6^3} \times \frac{2!}{6^2} = P(A_{24}) = \frac{12}{d}$$

$$P(A_1 A_4) = \frac{3!}{6^3} \times \frac{3!}{6^3} = \frac{6}{d}$$

$$P(A_1 A_2 A_3) = \frac{3!}{6^3} \times \frac{1}{6^2} = P(A_2 A_3 A_4) = \frac{6}{d}$$

$$P(A_1 A_2 A_4) = \frac{3!}{6^3} \times \frac{1}{6} \times \frac{2!}{6^2} = P(A_1 A_3 A_4) = \frac{2}{d}$$

$$P(A_1 A_2 A_3 A_4) = \frac{3!}{6^3} \times \frac{1}{6^3} = \frac{1}{d} .$$

So $p = 4 \times \dfrac{216}{d} - \left(3 \times \dfrac{36}{d} + 2 \times \dfrac{12}{d} + \dfrac{6}{d} \right) + \left(2 \times \dfrac{6}{d} + 2 \times \dfrac{2}{d} \right) - \dfrac{1}{d} = \dfrac{741}{7776} = \dfrac{247}{2592} .$

So the probability of no sweeps on six rolls is $1 - p = 1 - \dfrac{247}{2592} = \dfrac{2345}{2592} = \boxed{0.9047}$.

R Code for Problem 2

```
# (a)
trialfun=function(){
  n=2; v=sample(1:6,2,replace=T); cond=F; while(cond==F){
    n=n+1; v=c(v,sample(1:6,1)); ss=v[(n-2):n]
   if(   length( grep(1,ss)+grep(2,ss)+grep(3,ss) ) ==1) cond=T }
  n   }
set.seed(193); trialfun() # 92
date(); set.seed(284); J=10000; nv=rep(NA,J); for(j in 1:J) nv[j]=trialfun()
date()  # Took 32 secs
```

```
me=mean(nv); se=sd(nv); ci=me+c(-1,1)*qnorm(0.975)*se/sqrt(J)
c(me,se,ci)  # 43.88960 40.86393 43.08868 44.69052    OK

# (b) (Follows on from (a))
summary(nv)
#    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#    3.00   14.00   32.00   43.89   60.00  374.00
for(i in 1:400) fv[i]=length(nv[nv==i])
fv   #  [1]   0   0 295 225 2 .... 0   0   0   0
plot(1:400,fv);  sum(fv[seq(2,400,2)])/10000   # 0.502 OK

# (c)
trialfun2=function(n=38){
  v=sample(1:6,n,replace=T); y=0; for(i in 1:(n-2)){  ss=v[i:(i+2)]
      if(   length( grep(1,ss)+grep(2,ss)+grep(3,ss) ) ==1)  y=y+1 }
   y    }
set.seed(472); trialfun2() # 3 OK
date(); set.seed(224); J=10000; yv=rep(NA,J); for(j in 1:J) yv[j]=trialfun2()
date()  # Took 23 secs
me=mean(yv); se=sd(yv); ci=me+c(-1,1)*qnorm(0.975)*se/sqrt(J)
c(me,se,ci)  #    1.0020000 1.1256655 0.9799374 1.0240626   OK

# (d)
date(); set.seed(264); J=100000; yv=rep(NA,J); for(j in 1:J) yv[j]=trialfun2(n=6)
date()  # Took 25 secs
phat=length(yv[yv==0])/J
pci=phat+c(-1,1)*qnorm(0.975)*sqrt(phat*(1-phat)/J)
c(phat,pci) # 0.9061000 0.9042921 0.9079079   OK
```

Solution to Problem 3

(a) Let N be the total number of persons in the sample, and let Y be the number of persons in the sample with criplea. Then:

$$N \sim NegBin(w,q) \text{ with mean } EN = \frac{w}{q}, \text{ variance } VN = \frac{w(1-q)}{q^2}$$

$$\text{and density } f(n) = \binom{n-1}{w-1} q^w (1-q)^{n-w},\ n = w, w+1, w+2, \ldots$$

$$(Y \mid n) \sim Bin(n,p) \text{ with mean } E(Y \mid n) = np,\ \text{variance } V(Y \mid n) = np(1-p)$$

$$\text{and density } f(y \mid n) = \binom{n}{y} p^y (1-p)^{n-y},\ y = 0,1,\ldots,n.$$

So: $$EY = EE(Y \mid N) = E(Np) = pEN = \boxed{p\frac{w}{q}}$$

$$VY = EV(Y \mid N) + VE(Y \mid N) = E\{Np(1-p)\} + V\{Np\}$$

$$= p(1-p)EN + p^2 VN = p(1-p)\frac{w}{q} + p^2\frac{w(1-q)}{q^2} = \boxed{p\frac{w}{q}\left(1 + \frac{p}{q} - 2p\right)}.$$

With w = 200, q = 0.6 and p = 0.05 these formulae yield EY = $\boxed{16.67}$ and VY = $\boxed{16.39}$.

(b) $$f(n \mid y) = \frac{f(n)f(y \mid n)}{f(y)} = cf(n)f(y \mid n) \quad \text{where } c = \frac{1}{f(y)} \text{ does not depend on } n$$

$$= c\binom{n-1}{1-1} q^1 (1-q)^{n-1} \binom{n}{0} p^0 (1-p)^{n-0} \quad \text{since } w = 1 \text{ and } y = 0$$

$$= dt^{n-1},\ n = 1,2,3,\ldots$$

where d does not depend on n and where $t = (1-q)(1-p)$ = 0.637.

By considering a list of well-known discrete distributions, we see that N given $Y = y$ has a geometric distribution with parameter $k = 1 - t$. Thus:

$\boxed{(N \mid y) \sim Geo(k),\ \text{where } k = 0.363}$

$\boxed{f(n \mid y) = (1-k)^{n-1}k,\ n = 1,2,3,\ldots}$

$\boxed{E(N \mid y) = 1/k = 2.755}$.

(c) With q = 0.7 n = 5 and y = 2, the joint density of N and Y equals

$$f(n,y) = f(n)f(y\,|\,n) = \binom{5-1}{w-1}\left(\frac{7}{10}\right)^{w}\left(1-\frac{7}{10}\right)^{5-w} \times \binom{5}{2} p^2(1-p)^{5-2}.$$

So the likelihood function is

$$L(w,p) = \binom{4}{w-1} 7^w 3^{5-w} \times p^2(1-p)^3,\ 0 \le p \le 1,\ w = 1,\ldots,5.$$

We see that the MLE of p is $\boxed{2/5}$, because this value maximises $p^2(1-p)^3$.

Now, $\binom{4}{w-1} 7^w 3^{5-w}$ = 567, 5292, 18522, 28812, 16807 at w = 1, 2, 3, 4, 5,

respectively, with a maximum at w = 4. It follows that the MLE of w is $\boxed{4}$.

R Code for Problem 3

```
# (a)
w=200; q=0.6; p=0.05; me=p*w/q; va=me*(1-2*p+p/q)
c(me,va) # 16.66667 16.38889

# (b
q=0.35; p=0.02; t=(1-q)*(1-p); c(t,1-t,1/(1-t)) # 0.637000 0.363000 2.754821

# (c)
wv=1:5; choose(4, (1:5)-1) *7^wv * 3^(5-wv)  # 567  5292 18522 28812 16807
```

Solution to Problem 4

(a) We need to solve the equation $\mu = cEs$ for *c*, where

$$Es = E\left(\frac{\sigma}{\sqrt{n-1}}\sqrt{\frac{(n-1)s^2}{\sigma^2}}\right) = \frac{\mu}{\sqrt{n-1}}EU^{1/2} \quad \text{where } U = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1).$$

Now, $EU^{1/2} = \int_0^\infty u^{1/2}\frac{u^{\left(\frac{n-1}{2}\right)-1}e^{-\frac{u}{2}}}{\Gamma\left(\frac{n-1}{2}\right)2^{\left(\frac{n-1}{2}\right)}}du = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)2^{-\frac{1}{2}}}\int_0^\infty \frac{u^{\left(\frac{n}{2}\right)-1}e^{-\frac{u}{2}}}{\Gamma\left(\frac{n}{2}\right)2^{\left(\frac{n}{2}\right)}}du = \frac{\Gamma\left(\frac{n}{2}\right)\sqrt{2}}{\Gamma\left(\frac{n-1}{2}\right)}.$

Therefore $Es = \frac{\mu}{k}\sqrt{\frac{2}{n-1}}$ where $k = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$. So $\mu = cEs = c\frac{\mu}{k}\sqrt{\frac{2}{n-1}}$.

Thus $\boxed{c = k\sqrt{\frac{n-1}{2}}, \text{ where } k = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}}$.

For the case *n* = 6: $\Gamma\left(\frac{n-1}{2}\right) = \Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{3}{2}\times\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\sqrt{\pi}$

$$\Gamma\left(\frac{n}{2}\right) = \Gamma(3) = 2! = 2, \qquad k = \frac{3\sqrt{\pi}/4}{2} = \frac{3\sqrt{\pi}}{8} = 0.66467$$

$$c = k\sqrt{\frac{n-1}{2}} = \frac{3\sqrt{\pi}}{8}\sqrt{\frac{6-1}{2}} = \frac{3}{8}\sqrt{\frac{5\pi}{2}} = \boxed{1.050936}.$$

(b) Using results in (a), the variance of $\hat{\sigma}$ (as an unbiased estimator of $\sigma = \mu$) is

$$V\hat{\sigma} = c^2Vs = c^2\{Es^2 - (Es)^2\} = c^2\left\{\mu^2 - \left(\frac{\mu}{k}\sqrt{\frac{2}{n-1}}\right)^2\right\} = \mu^2c^2\left(1 - \frac{2}{k^2(n-1)}\right)$$

$$= \mu^2k^2\left(\frac{n-1}{2}\right)\left(1 - \frac{2}{k^2(n-1)}\right) = \mu^2\left\{\left(\frac{n-1}{2}\right)k^2 - 1\right\}.$$

Also $V\bar{y} = \dfrac{\sigma^2}{n} = \dfrac{\mu^2}{n}$. So, the efficiency of $\bar{y}$ relative to $\hat{\sigma}$ is

$$\boxed{r = \frac{V\hat{\sigma}}{V\bar{y}} = n\left(\left(\frac{n-1}{2}\right)k^2 - 1\right) \text{ where } k = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}}.$$

By results in (a) with n = 6 we get $r = 6\left(\left(\dfrac{6-1}{2}\right)\dfrac{9\pi}{64} - 1\right) = 6\left(\dfrac{45\pi}{128} - 1\right)$ = $\boxed{0.6268}$.

Note: The formula for k can be expressed in various other ways, for example using

the results that $\Gamma\left(\dfrac{m}{2}\right) = \left(\dfrac{m}{2} - 1\right)!$ if m is even and $\Gamma\left(\dfrac{m}{2}\right) = \dfrac{(m-2)!!\sqrt{\pi}}{2^{(m-1)/2}}$ if m is odd.

Here, !! denotes the double factorial function such that:

$m!! = m \times (m-2) \times ... \times 4 \times 2$ if m is even

$m!! = m \times (m-2) \times ... \times 3 \times 1$ if m is odd.

(c) The joint density is

$$f(y) = \prod_{i=1}^{n} \frac{1}{\mu\sqrt{2\pi}} \exp\left\{-\frac{1}{2\mu^2}(y_i - \mu)^2\right\} = \mu^{-n}(2\pi)^{-n/2} \exp\left\{-\frac{1}{2\mu^2}\sum_{i=1}^{n}(y_i - \mu)^2\right\},$$

and therefore the likelihood function is $L(\mu) = \mu^{-n} \exp\left\{-\dfrac{1}{2\mu^2}\sum_{i=1}^{n}(y_i - \mu)^2\right\}$.

So the log-likelihood function is $l(\mu) = -n\log\mu - \dfrac{1}{2}\mu^{-2}\sum_{i=1}^{n}(y_i - \mu)^2$.

Then, $l'(\mu) = -\dfrac{n}{\mu} - \dfrac{1}{2}\left\{\mu^{-2}2\sum_{i=1}^{n}(y_i - \mu)^1(-1) + (-2\mu^{-3})\sum_{i=1}^{n}(y_i - \mu)^2\right\}$

$$= -\frac{n}{\mu} + \left\{\mu^{-2}(n\bar{y} - n\mu) + \mu^{-3}(na - 2n\bar{y}\mu + n\mu^2)\right\} \text{ where } a = \frac{1}{n}\sum_{i=1}^{n} y_i^2$$

$$= \cancel{-\frac{n}{\mu}} + \left\{\mu^{-2}n\bar{y} - \mu^{-2}n\mu + \mu^{-3}na - \mu^{-3}2n\bar{y}\mu + \cancel{\mu^{-3}n\mu^2}\right\}$$

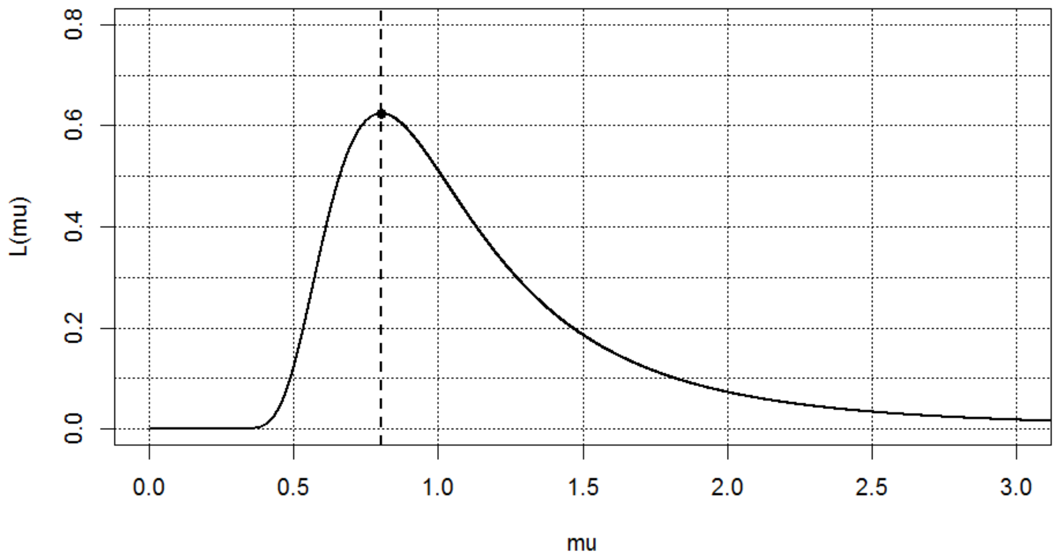
$$= -\mu^{-2}n\bar{y} - \mu^{-2}n\mu + \mu^{-3}na \qquad = -n\mu^{-3}\left(\mu^2 + \mu\bar{y} - a\right).$$

Setting $l'(\mu)$ to zero defined a quadratic equation whose solution yields the MLE,

$$\boxed{\hat{\mu} = \frac{-\bar{y} + \sqrt{\bar{y}^2 + 4a}}{2}, \text{ where } a = \frac{1}{n}\sum_{i=1}^{n} y_i^2} \quad \text{(since } \mu > 0 \text{)}.$$

If $(y_1,...,y_n)$ = (1.2,1.7, 0.1) then n = 3, $\bar{y}$ = 1, a = 1.44667 and $\hat{\mu}$ = $\boxed{0.8026}$.

The figure below shows $L(\mu)$ and shows the MLE with a vertical dashed line.



(d) The probability of a Type I error is

$$0.05 = P(S^2 > k) = P\left(\frac{(n-1)S^2}{\sigma^2} > \frac{(2-1)k}{3^2}\right) = P\left(\chi^2(1) > \frac{k}{9}\right)$$

$$= 2P\left(Z > \frac{\sqrt{k}}{3}\right) \quad \text{where } Z \sim N(0,1) \quad \text{since } Z^2 \sim \chi^2(1).$$

But $0.05 = 2P(Z > 1.96)$. So we equate $\frac{\sqrt{k}}{3} = 1.96$ to get $\sqrt{k} = 5.88$ and $k = \boxed{34.57}$.

Note: We could also get this value of k by looking up the chi-square tables with one degree of freedom, locating 3.84146 and multiplying this upper 0.05-quantile by 9.

The probability of a Type II error is

$$\beta(\mu) = P(S^2 < k) = P\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(2-1)k}{\mu^2}\right) = 1 - P\left(\chi^2(1) > \frac{k}{\mu^2}\right)$$

$$= 1 - 2P\left(Z > \frac{\sqrt{k}}{\mu}\right) \quad (\mu > 0,\ \mu \neq 3).$$

So the power function is $\boxed{Power(\mu) = 2P\left(Z > \frac{\sqrt{k}}{\mu}\right),\ \mu > 0}$.

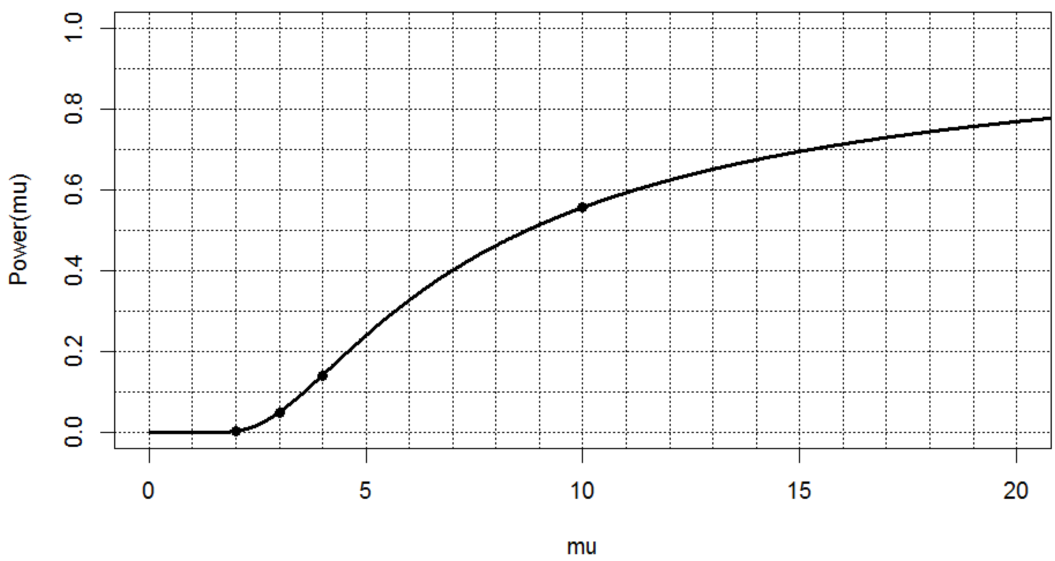
Thus: $Power(2) = 2P\left(Z > \frac{5.88}{2}\right) = 2P(Z > 2.94) = 2(0.0016) = \boxed{0.0032}$

$$Power(3) = 2P\left(Z > \frac{5.88}{3}\right) = \alpha = \boxed{0.0500}$$

$$Power(4) = 2P\left(Z > \frac{5.88}{4}\right) = 2P(Z > 1.47) = 2(0.0708) = \boxed{0.1416}$$

$$Power(10) = 2P\left(Z > \frac{5.88}{10}\right) = 2P(Z > 0.588) = 2(0.2776) = \boxed{0.5552}.$$

Note: These values were obtained using tables that are accurate to two decimals. The exact values correct to four decimals are 0.0033, 0.0500, 0.1416 and 0.5566. Below is a sketch of the power function with points marked at μ = 2, 3, 4 and 10.



R Code for Problem 4

```
# (a)
n=6; k=gamma((n-1)/2)/gamma(n/2); c=k*sqrt((n-1)/2)
c(k,c)  # 0.6646702 1.0509359

# (b)
n*(   ((n-1)/2)*k^2-1 ) # 0.626797
# Checking via Monte Carlo
mu=10;  J=10000; ybarvec=rep(NA,J); sighatvec=rep(NA,J); set.seed(294); for(j in 1:J){
    yv=rnorm(n,mu,mu); ybarvec[j]=mean(yv); sighatvec[j]=c*sd(yv) }
c(mean(ybarvec), mean(sighatvec)) # 10.01976 10.02061    OK
var(sighatvec)/var(ybarvec)  # 0.6226136   OK

# (c)
Lfun=function(mu,y){ n=length(y);    mu^(-n)*exp(-0.5*sum((y-mu)^2) /mu^2) }
X11(w=8,h=5); plot(c(0,3),c(0,0.8),type="n",xlab="mu",ylab="L(mu)");
    abline(h=seq(0,1,0.1),lty=3); abline(v=seq(0,5,0.5),lty=3)
muv=seq(0,10,0.001); m=length(muv); Lv=muv
y=c(1.2,1.7, 0.1); ybar=mean(y); a=mean(y^2); c(ybar,a) # 1.000000 1.446667
  for(i in 1:m) Lv[i]=Lfun(mu=muv[i],y=y); lines(muv,Lv,lty=1,lwd=2)
  mle=0.5*(-ybar+sqrt(ybar^2+4*a)); mle # 0.8025616
  points(mle, Lfun(mle,y), pch=16); abline(v=mle,lty=2,lwd=2)

# Another example
y=c(-0.3,-0.5,0.1); ybar=mean(y); a=mean(y^2); c(ybar,a) # -0.2333333  0.1166667
  for(i in 1:m) Lv[i]=Lfun(mu=muv[i],y=y); lines(muv,Lv,lty=3,lwd=2)
  mle=0.5*(-ybar+sqrt(ybar^2+4*a)); mle # 0.4776068
  points(mle, Lfun(mle,y), pch=16); abline(v=mle,lty=2,lwd=2)   # OK

# (d)
k=(3*qnorm(0.975))^2; k # 34.57313
2*(1-pnorm(sqrt(k)/c(2,3,4,10)))  # 0.003282695 0.050000000 0.141569070 0.556539545
```

```
X11(w=8,h=5); plot(c(0,20),c(0,1),type="n",xlab="mu",ylab="Power(mu)");
    abline(h=seq(0,1,0.1),lty=3); abline(v=seq(0,20,1),lty=3)
muvec=seq(0.01,21,0.01); betavec=2*(1-pnorm(sqrt(k)/muvec))
lines(muvec,betavec,lty=1,lwd=3)
points(c(2,3,4,10), 2*(1-pnorm(sqrt(k)/c(2,3,4,10))), pch=16)
```

Solution to Problem 5

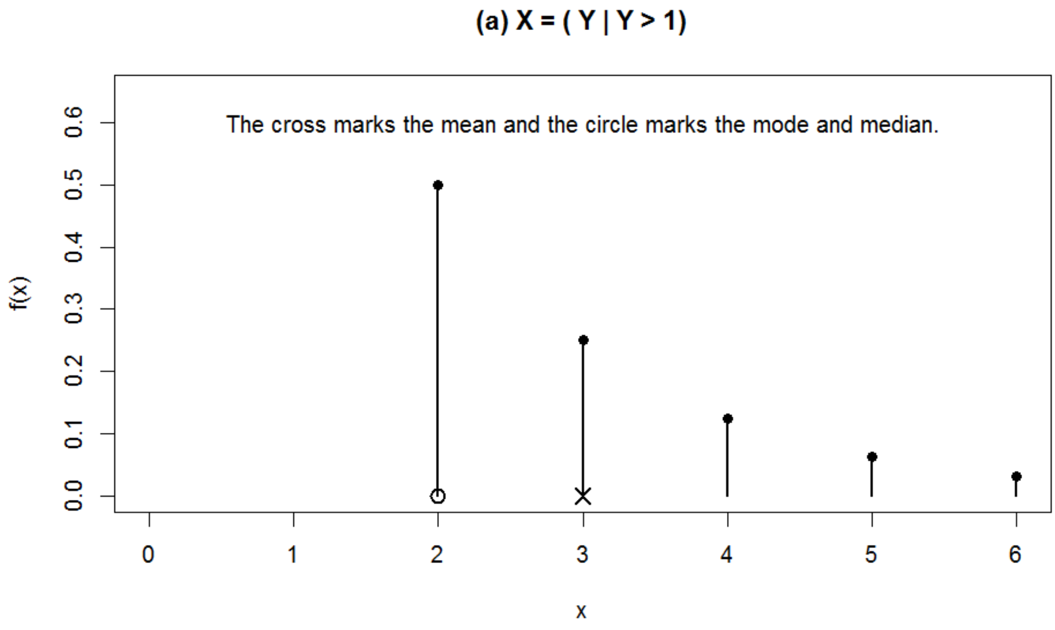
(a) $f_X(x) = \dfrac{f_Y(x)}{P(Y>1)}$, where $P(Y>1) = 1 - P(Y \le 1) = 1 - \dfrac{2}{3} - \dfrac{1}{3} \times \dfrac{1}{2} = \dfrac{1}{6}$.

So $\boxed{f_X(x) = 6 \times \dfrac{1}{3} \times \dfrac{1}{2^x} = \left(\dfrac{1}{2}\right)^{x-1},\ x = 2,3,4,\ldots}$ We see that $M = m = \boxed{2}$.

We may write $X = T + 1$, where $T \sim Geo(1/2)$ with mean $ET = \dfrac{1}{1/2} = 2$,

variance $VT = \dfrac{1-1/2}{(1/2)^2} = 2$ and second raw moment $ET^2 = VT + (ET)^2 = 6$.

Thus $\mu = EX = ET + 1 = 2 + 1 = \boxed{3}$, and $\sigma^2 = VX = VT = \boxed{2}$.



(b) $f_X(x) = \begin{Bmatrix} P(Y \le 1), & x = 0 \\ f_Y(x), & x = 2,3,4,\ldots \end{Bmatrix}$ where $P(Y \le 1) = \frac{2}{3} + \frac{1}{3} \times \frac{1}{2} = \frac{5}{6}$.

So $\boxed{f_X(x) = \begin{Bmatrix} 5/6, & x = 0 \\ \frac{1}{3}\left(\frac{1}{2}\right)^x, & x = 2,3,4,\ldots \end{Bmatrix}}$. We see that $M = m = \boxed{0}$.

Also: $\mu = 0 \times \frac{5}{6} + \frac{1}{3}\sum_{x=2}^{\infty} x\left(\frac{1}{2}\right)^x = \frac{1}{3}\sum_{t=1}^{\infty}(t+1)\left(\frac{1}{2}\right)^{t+1}$ where $t = x - 1$

$$= \frac{1}{6}\sum_{t=1}^{\infty}(t+1)\left(\frac{1}{2}\right)^t = \frac{1}{6}E(T+1) \quad \text{where } T \sim Geo(1/2)$$

$$= \frac{1}{6}(ET+1) = \frac{1}{6}(2+1) = \boxed{1/2}$$

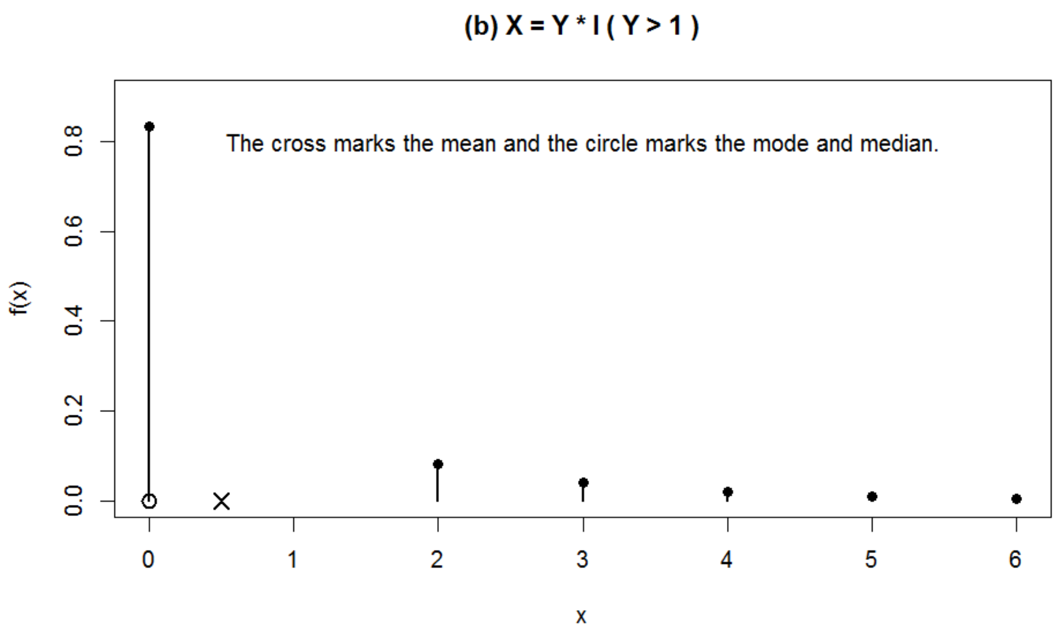
$$EX^2 = 0^2 \times \frac{5}{6} + \frac{1}{3}\sum_{x=2}^{\infty} x^2\left(\frac{1}{2}\right)^x = \frac{1}{3}\sum_{t=1}^{\infty}(t+1)^2\left(\frac{1}{2}\right)^{t+1}$$

$$= \frac{1}{6}\sum_{t=1}^{\infty}(t+1)^2\left(\frac{1}{2}\right)^t = \frac{1}{6}E\left\{(T+1)^2\right\} = \frac{1}{6}\left(ET^2 + 2ET + 1\right)$$

$$= \frac{1}{6}\left(VT + (ET)^2 + 2ET + 1\right)$$

$$= \frac{1}{6}\left(2 + 2^2 + 2 \times 2 + 1\right) = 11/6$$

$$VX = EX^2 - (EX)^2 = \frac{11}{6} - \left(\frac{1}{2}\right)^2 = \boxed{19/12}.$$



(c) $X = |Y-1|$ has pdf given by: $\quad f_X(0) = P(Y=1) = 1/6$

$$f_X(1) = P(Y=0) + P(Y=2) = \frac{1}{3} + \frac{1}{12} = 5/12$$

$$f_X(2) = P(Y=-1) + P(Y=3) = \frac{1}{6} + \frac{1}{24} = \frac{1}{2}\left(\frac{5}{12}\right), \text{ etc.}$$

We see that $\boxed{f_X(x) = \begin{cases} 1/6, & x = 0 \\ \frac{5}{12}\left(\frac{1}{2}\right)^{x-1}, & x = 1,2,3,\ldots \end{cases}}$ and $M = m = \boxed{1}$.

Also: $$\mu = EX = 0 \times \frac{1}{6} + 1 \times \frac{5}{12} + 2 \times \frac{5}{12}\left(\frac{1}{2}\right) + 3 \times \frac{5}{12}\left(\frac{1}{2}\right)^2 + \ldots$$

$$= \frac{5}{6}\left\{1\left(\frac{1}{2}\right) + 2\left(\frac{1}{2}\right)^2 + 3\left(\frac{1}{2}\right)^3 + \ldots\right\} = \frac{5}{12}ET \quad \text{where } T \sim Geo(1/2)$$

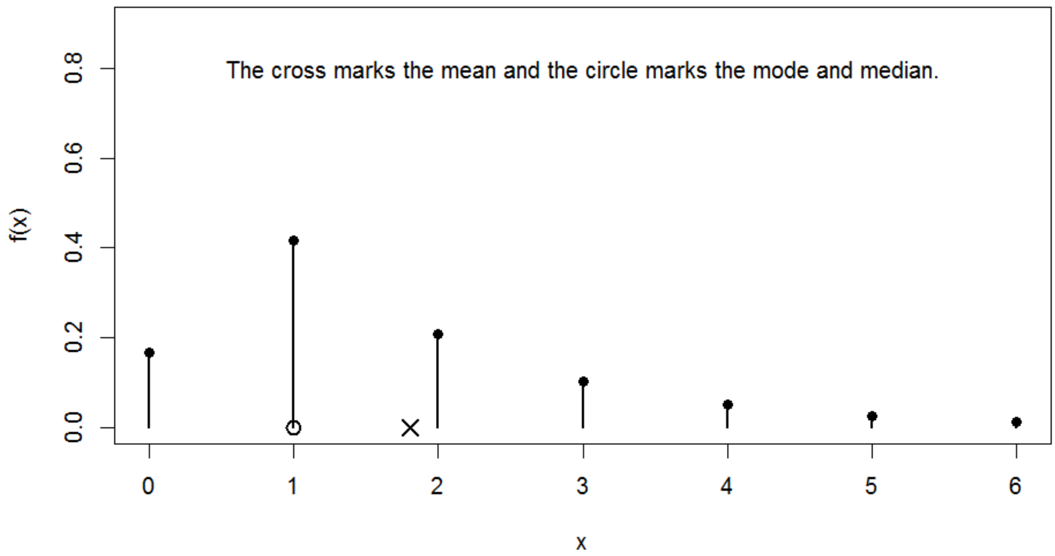
$$= \frac{5}{12} \times 2 = \boxed{5/6}$$

$$EX^2 = 0^2 \times \frac{1}{6} + 1^2 \times \frac{5}{12} + 2^2 \times \frac{5}{12}\left(\frac{1}{2}\right) + 3^2 \times \frac{5}{12}\left(\frac{1}{2}\right)^2 + \ldots$$

$$= \frac{5}{6}\left\{1^2\left(\frac{1}{2}\right) + 2^2\left(\frac{1}{2}\right)^2 + 3^2\left(\frac{1}{2}\right)^3 + \ldots\right\} = \frac{5}{12}ET^2 = \frac{5}{12} \times 6 = \frac{5}{2}$$

$$VX = EX^2 - (EX)^2 = \frac{5}{2} - \left(\frac{5}{6}\right)^2 = \boxed{65/36}.$$

(c) X = | Y - 1 |



Summary table (not required):

X	μ	M	m	σ^2
$(Y \mid Y > 1)$	3	2	2	2
$YI(Y > 1)$	1/2	0	0	19/12
$\lvert Y - 1 \rvert$	5/6	1	1	65/36

R Code for Problem 5

```
X11(w=8,h=5)

# (a)
xvec=2:6; plot(c(0,6),c(0,0.65),type="n",xlab="x",ylab="f(x)",
    main="(a) X = ( Y | Y > 1)")
fvec=(1/2)^(xvec-1); for(i in 1:length(xvec)){
     x=xvec[i]; lines(c(x,x),c(0,fvec[i]),lwd=2); points(x,fvec[i],pch=16)  }
points(c(2,3),c(0,0),pch=c(1,4),cex=1.5,lwd=2)
text(3,0.6,"The cross marks the mean and the circle marks the mode and median.")

# (b)
xvec=c(0,2:6); plot(c(0,6),c(0,0.9),type="n",xlab="x",ylab="f(x)",
    main="(b) X = Y * I ( Y > 1 )")
fvec=c(  5/6,   (1/3)*(1/2)^(2:6)  );  for(i in 1:length(xvec)){
     x=xvec[i]; lines(c(x,x),c(0,fvec[i]),lwd=2); points(x,fvec[i],pch=16)  }
points(c(0,1/2),c(0,0),pch=c(1,4),cex=1.5,lwd=2)
text(3,0.8,"The cross marks the mean and the circle marks the mode and median.")

# (c)
xvec=c(0:6); plot(c(0,6),c(0,0.9),type="n",xlab="x",ylab="f(x)",
    main="(c) X = | Y - 1 |")
fvec=c(  1/6,   (5/12)*(1/2)^(0:5)  );  for(i in 1:length(xvec)){
     x=xvec[i]; lines(c(x,x),c(0,fvec[i]),lwd=2); points(x,fvec[i],pch=16)  }
points(c(1,65/36),c(0,0),pch=c(1,4),cex=1.5,lwd=2)
text(3,0.8,"The cross marks the mean and the circle marks the mode and median.")
```

Solution to Problem 6

(a) Here: $\bar{x} = 2/3$, $\bar{y} = 8/3$,

$$S_{xx} = \sum x_i^2 - n\bar{x}^2 = 2 - 3(2/3)^2 = 2/3$$

$$S_{xy} = \sum x_i y_i - n\bar{x}\bar{y} = 7 - 3(2/3)(8/3) = 5/3,$$

$$b = S_{xy} / S_{xx} = \boxed{2.5}$$

$$a = \bar{y} - b\bar{x} = \boxed{1}$$

$$\hat{m} = a + 0.5b = \boxed{2.25}.$$

Next, $V\hat{m} = 1\left(1 + \frac{1}{3} + \frac{(0.5 - 2/3)^2}{2/3}\right) = \frac{11}{8}$, and so a 95% prediction interval for m is

$$\left(\hat{m} \pm 1.96\sqrt{V\hat{m}}\right) = \left(2.25 \pm 1.96\sqrt{11/8}\right) = (2.25 \pm 2.298) = \boxed{(-0.048, 4.548)}.$$

(b) The sum of squares for error is $SSE = \sum(y_i - au_i - bx_i)^2$. This has derivatives:

$$\frac{\partial SSE}{\partial a} = \sum 2(y_i - au_i - bx_i)^1(-u_i) = -2\left\{T_{yu} - aT_{uu} - bT_{xu}\right\}$$

where $T_{yu} = \sum y_i u_i$, $T_{uu} = \sum u_i u_i = \sum u_i^2$, etc. ($T$ stands for "Total")

$$\frac{\partial SSE}{\partial b} = \sum 2(y_i - au_i - bx_i)^1(-x_i) = -2\left\{T_{yx} - aT_{ux} - bT_{xx}\right\}.$$

Setting these derivatives to zero yields $b = \frac{T_{yu} - aT_{uu}}{T_{xu}}$ and $\boxed{b = \frac{T_{yx} - aT_{ux}}{T_{xx}}}$.

Equating these two expressions for b, we obtain $\boxed{a = \frac{T_{xx}T_{uy} - T_{ux}T_{xy}}{T_{uu}T_{xx} - T_{ux}^2}}$.

For the given data, we get $a = \frac{2\times 9 - 2\times 7}{5\times 2 - 2^2} = \boxed{2/3}$ and $b = \frac{7 - (2/3)2}{2} = \boxed{17/6}$.