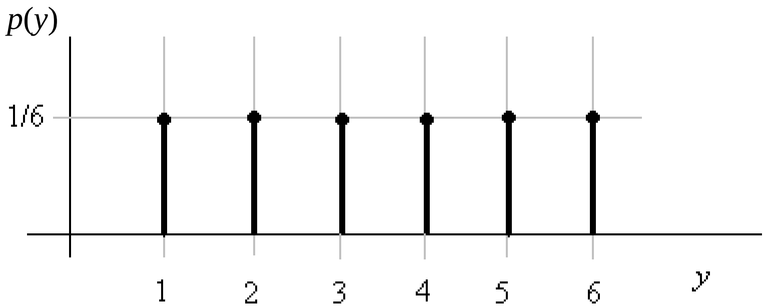
# STAT2001 Tutorial 5 Solutions

**(a)** $p(y) = 1/6, \quad y = 1, \ldots, 6$.




**(b)** $EY = \sum_{y=1}^{6} y \frac{1}{6} = \frac{1}{6}(1+2+...+6) = \frac{7}{2}$.

This makes sense because 3.5 is halfway between 1 and 6.

$EY^2 = \sum_{y=1}^{6} y^2 \frac{1}{6} = \frac{1}{6}(1^2+2^2+...+6^2) = \frac{91}{6}$.

$VarY = EY^2 - (EY)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = 2.9167$.     $SD(Y) = \sqrt{2.9167} = 1.708$.

*A similar problem:* A number $Y$ is randomly chosen from 1,2,...,100.
Find $EY$ and $VarY$.

$$EY = \sum_{y=1}^{100} y \frac{1}{100} = \frac{1}{100}(1+2+...+100) = \frac{1}{100}\frac{100(101)}{2} = 50.5.$$

We have here used the fact that $1 + 2 + ... + n = n(n + 1)/2$.

*Proof:* Let $s = 1 + 2 + ... + n$. Then

$$\begin{aligned} 2s &= 1 + 2 + \ldots + (n-1) + n \\ &\quad + n + (n-1) + \ldots + 2 + 1 \\ &= (n+1) + (n+1) + \ldots + (n+1) + (n+1). \end{aligned}$$

We see that $2s = n(n + 1)$. Therefore $s = n(n + 1)/2$.

Also, $EY^2 = \sum_{y=1}^{100} y^2 \frac{1}{100} = \frac{1}{100}(1^2+2^2+...+100^2) = \frac{1}{100}\frac{100(101)201}{6} = 3383.5$.

We have here used the fact that $1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$.

Hence $VarY = 3383.5 - 50.5^2 = 833.25$.

NB: The results used here are also given in Appendix 1 of the text (7th edition, p836).

**Problem 2**

**(a)** Let $Y$ = number on die and $G$ = Kate's net gain (in \$'s).
Then $G = 3Y - 11$.
So $EG = 3EY - 11 = 3(3.5) - 11 = -0.5$ (by Problem 1).
That is, Kate can expect to lose 50 cents overall. The game is not fair.

**(b)** $VarG = Var(3Y - 11) = 3^2 VarY = 9 \times 2.9167 = 26.25$.

**Problem 3**

**(a)**
$$m(t) = Ee^{Yt} = \sum_{y=1}^{\infty} e^{yt} q^{y-1} p \quad \text{where } q = 1 - p$$
$$= pe^t \sum_{y=1}^{\infty} (qe^t)^{y-1} = pe^t \sum_{x=0}^{\infty} (qe^t)^x \quad \text{where } x = y - 1$$
$$= \frac{pe^t}{1 - qe^t} = \frac{pe^t}{1 - (1-p)e^t}.$$

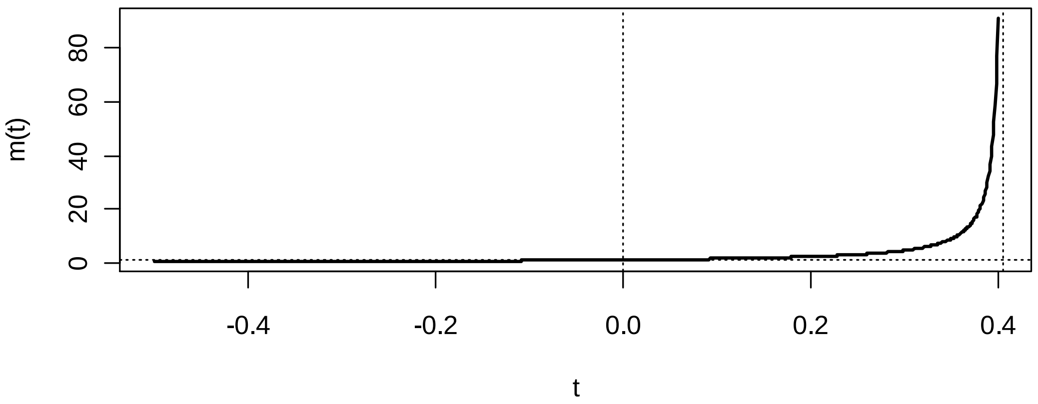
Note: $m(t)$ is defined only for $t$ such that $-1 < qe^t < 1$, ie $t < \log(q^{-1}) = -\log q$.
$m(t)$ asymptotes to infinity as $t$ approaches $-\log q$ from the left.

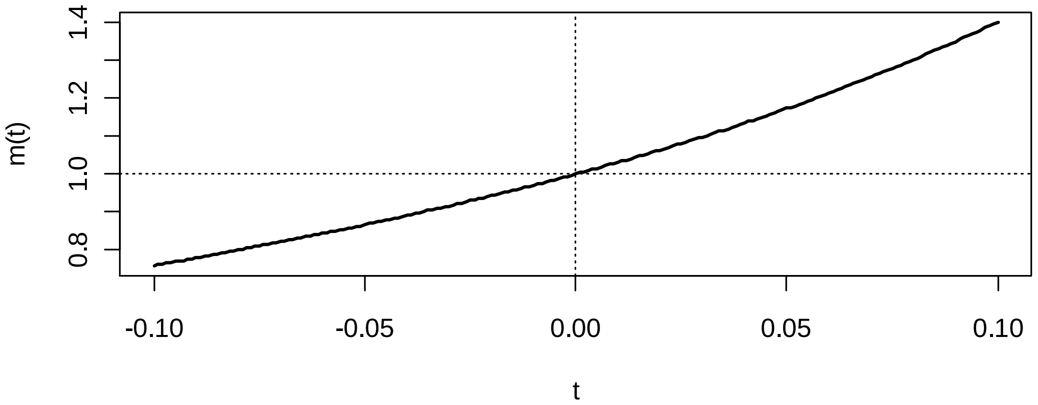
*Example*

The following figure shows the mgf of $Y$ when $p = 1/3$, on two different scales.

The top plot shows how $m(t) \to \infty$ as $t \to -\log q = -\log(2/3) = 0.4055$.

The bottom plot shows that $m(0) = 1$, and also that $m'(0) = EY = 1/p = 3$, or in other words, that the slope of the mgf at $t = 0$ is the mean of $Y$ (see part(b)).




**R Code (non-assessable)**

```
p = 1/3; q = 1-p; log(1/q) # 0.4054651
par(mfrow=c(2,1));

tv=seq(-0.5, 0.4,0.001);
mv = p*exp(tv)/(1-q*exp(tv))
plot(tv,mv,type="l",xlab="t",ylab="m(t)",lwd=2)
abline(v=0,h=1,lty=3); abline(v= log(1/q), lty=3)

tv=seq(-0.1, 0.1,0.001);
mv = p*exp(tv)/(1-q*exp(tv))
plot(tv,mv,type="l",xlab="t",ylab="m(t)",lwd=2)
abline(v=0,h=1,lty=3)
```

**(b)** $m(t) = pe^t(1-qe^t)^{-1}$.

So $m'(t) = p\{e^t(-1)(1-qe^t)^{-2}(-qe^t) + (1-qe^t)^{-1}e^t\}$

$$= \frac{q}{p}\left\{pe^t(1-qe^t)^{-1}\right\}^2 + \left\{pe^t(1-qe^t)^{-1}\right\}$$

$$= \frac{q}{p}m(t)^2 + m(t).$$

So $\mu = m'(0) = \frac{q}{p}m(0)^2 + m(0) = \frac{q}{p}1^2 + 1 = \frac{q+p}{p} = \frac{1}{p}$.

Next, $m''(t) = \frac{d}{dt}\left(\frac{q}{p}m(t)^2 + m(t)\right) = \frac{q}{p}2m(t)m'(t) + m'(t)$.

So $\mu_2' = m''(0) = \frac{q}{p}2(1)\frac{1}{p} + \frac{1}{p} = \frac{2q+p}{p^2} = \frac{1+q}{p^2}$.

Hence $\sigma^2 = \mu_2' - \mu^2 = \frac{1+q}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{q}{p^2} = \frac{1-p}{p^2}$.

**Another way to find $\mu$ (Method 2: Direct summation)**

$\mu = \sum_{y=1}^{\infty} yq^{y-1}p = \frac{p}{q}s$, where $s = \sum_{y=1}^{\infty} yq^y = 1q^1 + 2q^2 + 3q^3 + \ldots$

Now $qs = q^2 + 2q^3 + 3q^4 + \ldots$

$$= (2q^2 + 3q^3 + 4q^4 + \ldots) - (q^2 + q^3 + \ldots)$$

$$= (s-q) - \{(1 + q + q^2 + q^3 + \ldots) - 1 - q\}.$$

Therefore $qs = (s-q) - \left(\frac{1}{1-q} - 1 - q\right)$.

Solving this equation, we find that $s = \frac{q}{p^2}$.

It follows that $\mu = \frac{p}{q}s = \frac{p}{q}\frac{q}{p^2} = \frac{1}{p}$.

**Another way to find $\mu$ (Method 3: Differential calculus)**

$$\mu = p\sum_{y=1}^{\infty} yq^{y-1} = p\sum_{y=1}^{\infty} \frac{dq^y}{dq} = p\frac{d}{dq}\sum_{y=1}^{\infty} q^y = p\frac{d}{dq}\left(\frac{1}{1-q}-1\right) = p\frac{1}{(1-q)^2} = \frac{1}{p}.$$

**Yet another way to find $\mu$ (Method 4: First step analysis)**

Recall the law of total probability (LTP):

$$P(A) = P(B)P(A \mid B) + P(\bar{B})P(A \mid \bar{B}).$$

There also exists a similar law called the *law of total expectation (LTE)*:

$$EY = P(B)E(Y \mid B) + P(\bar{B})E(Y \mid \bar{B}).$$

Let $B$ be the event that a success occurs on the first trial.

Then by the LTE,

$$\mu = p\times 1 + (1-p)\times(\mu + 1). \qquad (*)$$

(If a success *doesn't* occur on the first trial, the expected number of trials *from that point on* is the same as it was in the beginning, namely $\mu$. Therefore the expected *total* number of trials must equal $\mu$ plus 1.)

Rearranging (*) leads to $\mu = 1/p$, as before.

**Problem 4**

**(a)** Let $Y$ be the number of rolls. Then $Y \sim \text{Geo}(1/6)$.

Hence by Problem 3: $EY = \dfrac{1}{1/6} = 6$ and $VarY = \dfrac{5/6}{(1/6)^2} = 30$.

So $EY^2 = VarY + (EY)^2 = 30 + 6^2 = 66$.

Next let $G$ = Kate's gain (in $'s). Then $G = 2Y^2 - 100$.

Therefore $EG = 2EY^2 - 100 = 2(66) - 100 = \$32$.

**(b)** $P(G > 0) = P(2Y^2 - 100 > 0) = P(Y^2 > 50) = P(Y > 7.07) = P(Y \geq 8)$

$= P(\text{No 6's come up on the first 7 rolls}) = (5/6)^7 = 0.279$.

*Additional question*

What is the probability that Kate's net gain will be *negative*?

Observe that $P(G = 0) = P(2Y^2 - 100 = 0) = P(Y = 7.07) = 0$.

Therefore $P(G < 0) = 1 - P(G = 0) - P(G > 0) = 1 - 0 - 0.279 = 0.721$.

*Discussion*

It may seem counterintuitive that Kate can expect to gain a positive amount (\$32) and yet will most likely lose money (with probability 72.1%).

However, such situations are not uncommon.

For example, consider the random variable $X$ whose pdf is $p(x) = \begin{cases} 0.1, & x = 100 \\ 0.9, & x = -1. \end{cases}$

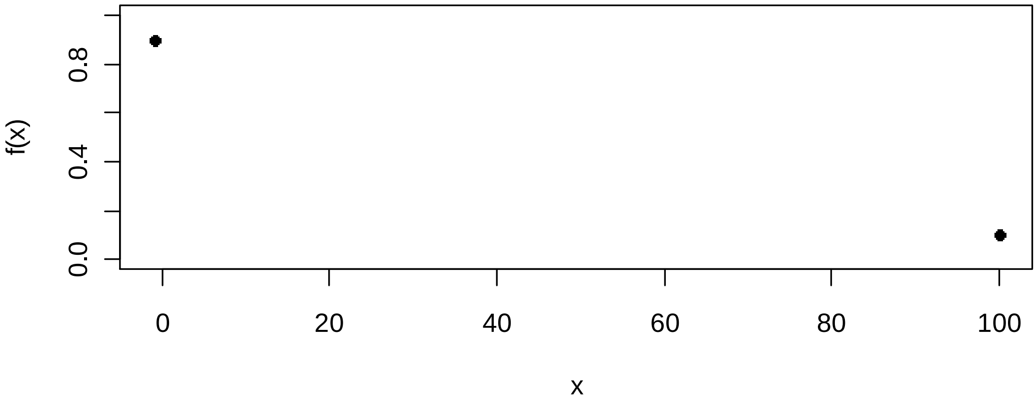
Then: $EX = 100(0.1) + (-1)(0.9) = 9.1$
$P(X < 0) = 0.9$.

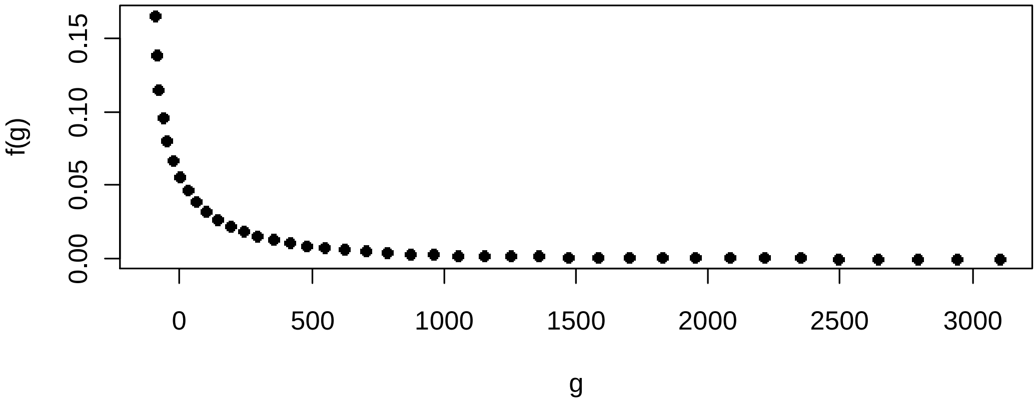
Thus $X$ has a positive expectation (9.1), and yet at the same time a high probability of being negative (0.9).

Kate's situation is similar to this, only more complicated.

The following figure shows the pdf of $X$ here and the pdf of $G$ (Kate's gain).

(Note: The theory behind the derivation of $G$'s pdf is a topic which will be covered in Chapter 6.)




**R Code  (non-assessable)**

```
par(mfrow=c(2,1))
plot(c(-1,100),c(0.9,0.1),pch=16,xlab="x",ylab="f(x)",ylim=c(0,1))

Yv=1:200; pYv=dgeom(Yv-1,1/6) # NB: R uses a different defn of the geom. dsn
c(sum(pYv), sum(Yv*pYv), sum(Yv^2*pYv)) # 1 6 66 (correct)
2*66 - 100  # 32
Gv=2*Yv^2-100;  sum(Gv*pYv)  # 32
plot(Gv[1:40],pYv[1:40],pch=16,xlab="g",ylab="f(g)")
sum(pYv[Gv<0])  # 0.7209184
```