1. (a) From the basic rule of counting, there are

$$\binom{4}{2} \times \binom{4}{2} = 6 \times 6 = 36$$

possible outcomes from the experiment of choosing 2 puppies from each litter. Let

- C_1, B_1, Y_1, and Y_2 denote the four puppies in Litter 1
- B_2, B_3, Y_3, Y_4 denote the four puppies in Litter 2.

I have listed the 36 outcomes below, along with the realized value of Y for each outcome. Each outcome is equally likely.

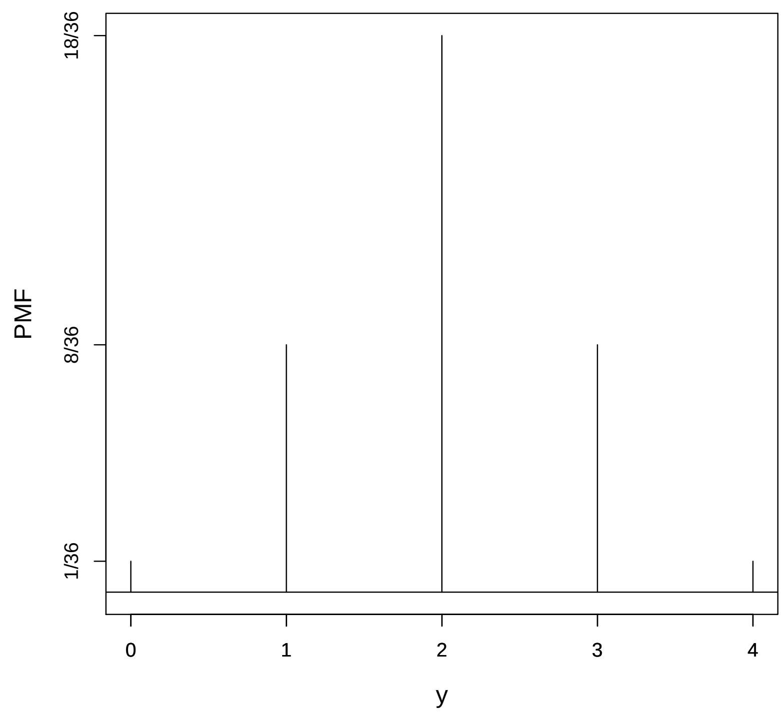
Outcome	y	probability	Outcome	y	probability
$(C_1, B_1 \| B_2, B_3)$	0	$\frac{1}{36}$	$(C_1, Y_1 \| B_2, B_3)$	1	$\frac{1}{36}$
$(C_1, B_1 \| B_2, Y_3)$	1	$\frac{1}{36}$	$(C_1, Y_1 \| B_2, Y_3)$	2	$\frac{1}{36}$
$(C_1, B_1 \| B_2, Y_4)$	1	$\frac{1}{36}$	$(C_1, Y_1 \| B_2, Y_4)$	2	$\frac{1}{36}$
$(C_1, B_1 \| B_3, Y_3)$	1	$\frac{1}{36}$	$(C_1, Y_1 \| B_3, Y_3)$	2	$\frac{1}{36}$
$(C_1, B_1 \| B_3, Y_4)$	1	$\frac{1}{36}$	$(C_1, Y_1 \| B_3, Y_4)$	2	$\frac{1}{36}$
$(C_1, B_1 \| Y_3, Y_4)$	2	$\frac{1}{36}$	$(C_1, Y_1 \| Y_3, Y_4)$	3	$\frac{1}{36}$
$(C_1, Y_2 \| B_2, B_3)$	1	$\frac{1}{36}$	$(B_1, Y_1 \| B_2, B_3)$	1	$\frac{1}{36}$
$(C_1, Y_2 \| B_2, Y_3)$	2	$\frac{1}{36}$	$(B_1, Y_1 \| B_2, Y_3)$	2	$\frac{1}{36}$
$(C_1, Y_2 \| B_2, Y_4)$	2	$\frac{1}{36}$	$(B_1, Y_1 \| B_2, Y_4)$	2	$\frac{1}{36}$
$(C_1, Y_2 \| B_3, Y_3)$	2	$\frac{1}{36}$	$(B_1, Y_1 \| B_3, Y_3)$	2	$\frac{1}{36}$
$(C_1, Y_2 \| B_3, Y_4)$	2	$\frac{1}{36}$	$(B_1, Y_1 \| B_3, Y_4)$	2	$\frac{1}{36}$
$(C_1, Y_2 \| Y_3, Y_4)$	3	$\frac{1}{36}$	$(B_1, Y_1 \| Y_3, Y_4)$	3	$\frac{1}{36}$
$(B_1, Y_2 \| B_2, B_3)$	1	$\frac{1}{36}$	$(Y_1, Y_2 \| B_2, B_3)$	2	$\frac{1}{36}$
$(B_1, Y_2 \| B_2, Y_3)$	2	$\frac{1}{36}$	$(Y_1, Y_2 \| B_2, Y_3)$	3	$\frac{1}{36}$
$(B_1, Y_2 \| B_2, Y_4)$	2	$\frac{1}{36}$	$(Y_1, Y_2 \| B_2, Y_4)$	3	$\frac{1}{36}$
$(B_1, Y_2 \| B_3, Y_3)$	2	$\frac{1}{36}$	$(Y_1, Y_2 \| B_3, Y_3)$	3	$\frac{1}{36}$
$(B_1, Y_2 \| B_3, Y_4)$	2	$\frac{1}{36}$	$(Y_1, Y_2 \| B_3, Y_4)$	3	$\frac{1}{36}$
$(B_1, Y_2 \| Y_3, Y_4)$	3	$\frac{1}{36}$	$(Y_1, Y_2 \| Y_3, Y_4)$	4	$\frac{1}{36}$

Therefore, the probability mass function of Y is

y	0	1	2	3	4
$p_Y(y)$	1/36	8/36	18/36	8/36	1/36

A graph of this pmf is at the top of the next page.
(b) The moment generating function of Y is

$$\begin{aligned} m_Y(t) = E(e^{tY}) &= \sum_{y=0}^{4} e^{ty} p_Y(y) \\ &= e^{t(0)}\left(\frac{1}{36}\right) + e^{t(1)}\left(\frac{8}{36}\right) + e^{t(2)}\left(\frac{18}{36}\right) + e^{t(3)}\left(\frac{8}{36}\right) + e^{t(4)}\left(\frac{1}{36}\right) \\ &= \frac{1}{36} + \left(\frac{8}{36}\right)e^{t} + \left(\frac{18}{36}\right)e^{2t} + \left(\frac{8}{36}\right)e^{3t} + \left(\frac{1}{36}\right)e^{4t}. \end{aligned}$$



2. (a) We know $0 \leq p_Y(y) \leq 1$ for all y. Therefore, $0 \leq p \leq 1$ and

$$0 \leq 1 - 2p \leq 1 \iff -1 \leq -2p \leq 0 \iff \frac{1}{2} \geq p \geq 0.$$

We must have $0 \leq p \leq 1/2$. Note that if $p = 0$, then $p_Y(0) = 1$; i.e., Y has a degenerate distribution at $y = 0$. If $p = 1/2$, then $p_Y(y) = 0$ and $y = \pm 2$ each have probability $1/2$.
(b) The mean of Y is

$$E(Y) = \sum_{y \in R} y p_Y(y) = -2(p) + 0(1 - 2p) + 2(p) = -2p + 2p = 0.$$

The second moment of Y is

$$E(Y^2) = \sum_{y \in R} y^2 p_Y(y) = (-2)^2(p) + 0^2(1 - 2p) + 2^2(p) = 4p + 4p = 8p.$$

Therefore,

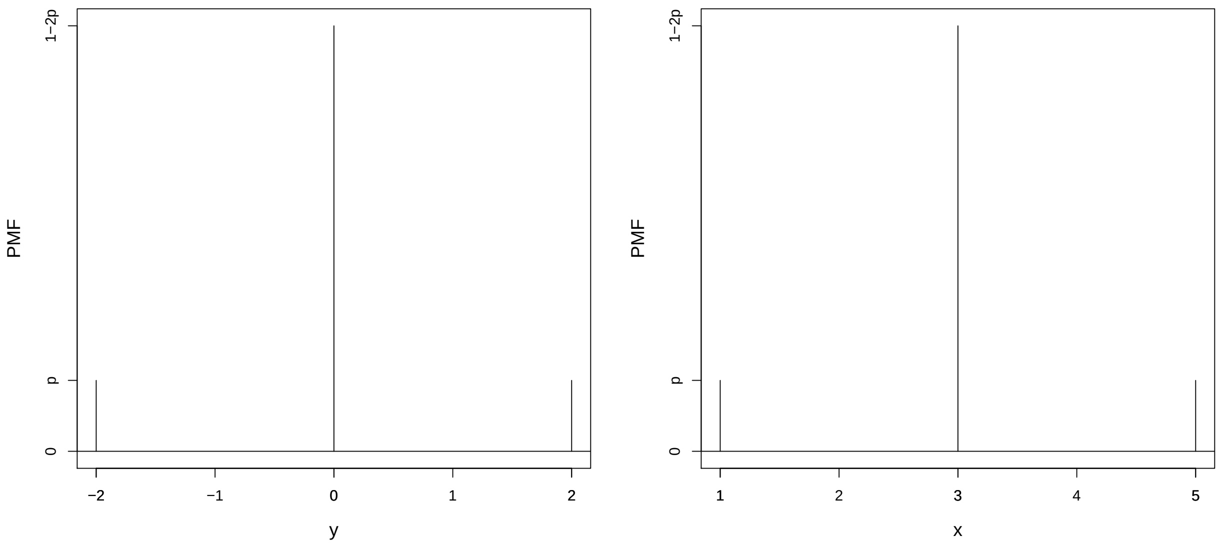
$$V(Y) = E(Y^2) - [E(Y)]^2 = 8p - (0)^2 = 8p,$$

as claimed.
(c) We have determined $\mu = E(Y) = 0$. Therefore, the skewness associated with Y is

$$\xi = \frac{E[(Y - \mu)^3]}{\sigma^3} = \frac{E(Y^3)}{(\sqrt{8p})^3}.$$

However,

$$E(Y^3) = \sum_{y \in R} y^3 p_Y(y) = (-2)^3(p) + 0^3(1 - 2p) + 2^3(p) = -8p + 8p = 0.$$



Therefore, $\xi = 0$; i.e., the distribution of Y is symmetric about $\mu = 0$. The kurtosis associated with Y is

$$\kappa = \frac{E[(Y-\mu)^4]}{\sigma^4} = \frac{E(Y^4)}{(\sqrt{8p})^4}.$$

Note that

$$E(Y^4) = \sum_{y \in R} y^4 p_Y(y) = (-2)^4(p) + 0^4(1-2p) + 2^4(p) = 16p + 16p = 32p.$$

Therefore,

$$\kappa = \frac{32p}{64p^2} = \frac{1}{2p}.$$

Values of p close to zero will inflate the kurtosis. This makes sense; for values of p close to zero, $p_Y(0)$ will be very close to one and the distribution will be more peaked relative to the other two values $y = \pm 2$.
(d) The pmf of Y is shown at the top of the page (left). Adding 3 to each value of Y will shift the distribution of Y three units to the right (right). Clearly,

$$E(X) = E(Y+3) = E(Y) + 3 = 0 + 3 = 3.$$

All the other values will remain the same.

- The variance σ^2 measures spread, which is unaffected by the additive shift; i.e., $V(X) = V(Y)$.
- The additive shift does not affect the skewness ξ or the kurtosis κ. The pmf of X (above, right) is now symmetric about $E(X) = 3$ and the amount of peakedness in the distribution of X is the same as that in the distribution of Y.

3. (a) We think of each customer as a "trial." In the context of this problem, the three Bernoulli trial assumptions are:

- Each customer either has an expired license ("success") or a license that is not expired.
- The probability $p = 0.30$ of having an expired license is the same for every customer.
- The expired/not expired license statuses of the customers are mutually independent.

(b) We have a fixed number of customers here ($n = 8$). Let

$$Y = \text{ number of customers who have an expired license (out of } n = 8).$$

Under the Bernoulli trial assumptions in part (a), we have $Y \sim b(n = 8, p = 0.30)$. We want to calculate $P(Y \leq 2)$. We have

$$\begin{aligned} P(Y \leq 2) &= P(Y = 0) + P(Y = 1) + P(Y = 2) \\ &= \binom{8}{0}(0.30)^0(0.70)^8 + \binom{8}{1}(0.30)^1(0.70)^7 + \binom{8}{2}(0.30)^2(0.70)^6 \\ &\approx 0.058 + 0.198 + 0.296 = 0.552. \end{aligned}$$

Checking my work in R,

```
> pbinom(2,8,0.30)
[1] 0.5517738
```

(c) Under the Bernoulli trial assumptions in part (a), $W \sim \text{geom}(p = 0.30)$.

Solution 1: By the complement rule (for conditional probabilities), we have

$$P(W \leq 5 | W > 2) = 1 - P(W > 5 | W > 2).$$

From Problem 3.71 (HW5), we have

$$P(W > 5 | W > 2) = P(W > 3) = (0.7)^3 = 0.343.$$

Therefore,

$$P(W \leq 5 | W > 2) = 1 - 0.343 = 0.657.$$

Solution 2: The solution above makes use of the *memoryless property* which is associated with the geometric distribution; see Problem 3.71. We could also do this problem from first principles by defining the events $A = \{W \leq 5\}$ and $B = \{W > 2\}$. We want to calculate $P(A|B)$. From the definition of conditional probability, we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(W \leq 5 | W > 2).$$

Note that

$$A \cap B = \{W \leq 5\} \cap \{W > 2\} = \{W = 3, 4, 5\}$$

Therefore,

$$\begin{aligned} P(A \cap B) &= P(W=3) + P(W=4) + P(W=5) \\ &= (0.70)^{3-1}(0.30) + (0.70)^{4-1}(0.30) + (0.70)^{5-1}(0.30) \approx 0.322. \end{aligned}$$

In addition,

$$\begin{aligned} P(B) = P(W>2) &= 1 - P(W=1) - P(W=2) \\ &= 1 - [(0.70)^{1-1}(0.30) + (0.70)^{2-1}(0.30)] = 0.49. \end{aligned}$$

Finally,

$$P(W \leq 5 | W > 2) = \frac{P(A \cap B)}{P(B)} \approx \frac{0.322}{0.49} = 0.657.$$

(c) Let X denote the number of customers the employee will help until she finds the 3rd customer with an expired license (i.e., the 3rd "success."). Under the Bernoulli trial assumptions in part (a), $X \sim \text{nib}(r=3, p=0.30)$. The pmf of X is

$$p_X(x) = \begin{cases} \binom{x-1}{2}(0.30)^3(0.70)^{x-3}, & x = 3,4,5,... \\ 0, & \text{otherwise.} \end{cases}$$

4. (a) The mean and variance of Y are each equal to $\lambda = 10$. Therefore,

$$\mu + 2\sigma = 10 + 2\sqrt{10} \approx 16.32$$

and $P(Y > \mu + 2\sigma) = P(Y \geq 17)$. From the complement rule, we have

$$P(Y \geq 17) = 1 - P(Y \leq 16) = 1 - \underbrace{\sum_{y=0}^{16} \frac{10^y e^{-10}}{y!}}_{\texttt{ppois(16,10)}} \approx 0.027.$$

The R function `ppois` calculates $P(Y \leq 16)$ for us. We have

```
> 1-ppois(16,10)
[1] 0.02704161
```

(b) First note that

$$\begin{aligned} Q = 3Y^4 + (2Y-1)^2 + 10000 &= 3Y^4 + (4Y^2 - 4Y + 1) + 10000 \\ &= 3Y^4 + 4Y^2 - 4Y + 10001. \end{aligned}$$

Therefore,

$$\begin{aligned} E(Q) = E(3Y^4 + 4Y^2 - 4Y + 10001) &= E(3Y^4) + E(4Y^2) + E(-4Y) + E(10001) \\ &= 3E(Y^4) + 4E(Y^2) - 4E(Y) + 10001. \end{aligned}$$

We know that $E(Y) = 10$. In addition,

$$10 = V(Y) = E(Y^2) - [E(Y)]^2 = E(Y^2) - 10^2 \implies E(Y^2) = 10 + 10^2 = 110.$$

How do we derive $E(Y^4)$? The only way I know to do it is to use the mgf; i.e.,

$$E(Y^4) = \frac{d^4}{dt^4} m_Y(t)\Big|_{t=0}.$$

We need to take 4 derivatives of $m_Y(t) = e^{10(e^t-1)}$. Here we go!

$$\begin{aligned}
\frac{d}{dt} m_Y(t) &= 10e^t \times e^{10(e^t-1)} = 10e^{10(e^t-1)+t} \\
\frac{d^2}{dt^2} m_Y(t) &= 10(10e^t+1)e^{10(e^t-1)+t} \\
\frac{d^3}{dt^3} m_Y(t) &= 10\left[(10e^t)e^{10(e^t-1)+t} + (10e^t+1)(10e^t+1)e^{10(e^t-1)+t}\right] \\
&= 10e^{10(e^t-1)+t}\left[10e^t + (10e^t+1)^2\right].
\end{aligned}$$

Finally, we have

$$\frac{d^4}{dt^4} m_Y(t) = 10(10e^t+1)e^{10(e^t-1)+t}\left[10e^t + (10e^t+1)^2\right] + 10e^{10(e^t-1)+t}[10e^t + 2(10e^t+1)10e^t].$$

Therefore,

$$\begin{aligned}
E(Y^4) &= 10(10e^0+1)e^{10(e^0-1)+0}\left[10e^0 + (10e^0+1)^2\right] + 10e^{10(e^0-1)+0}[10e^0 + 2(10e^0+1)10e^0] \\
&= 10(11)e^0\left[10 + (10+1)^2\right] + 10e^0[10 + 2(10+1)10] \\
&= 10(11)(131) + 10(10+220) \\
&= 16710.
\end{aligned}$$

Finally,

$$E(Q) = 3E(Y^4) + 4E(Y^2) - 4E(Y) + 10001 = 3(16710) + 4(110) - 4(10) + 10001 = 60531.$$

Therefore, the expected daily cost of quarantining individuals who test positively is \$60,531.
(c) I have plotted the Poisson(10) pmf at the top of the next page, and I have used dark circle plotting symbols to indicate the 10 daily count observations during September 26-October 5. It is certainly possible this is an appropriate model given where the observations fall (although we would expect more to fall in the center−around 10). The "69" count is not an expected observation from this distribution, however. This observation could be the result of testing the specimens which may have backed up over the weekend (October 5 was a Monday).

5. (a) It suffices to show

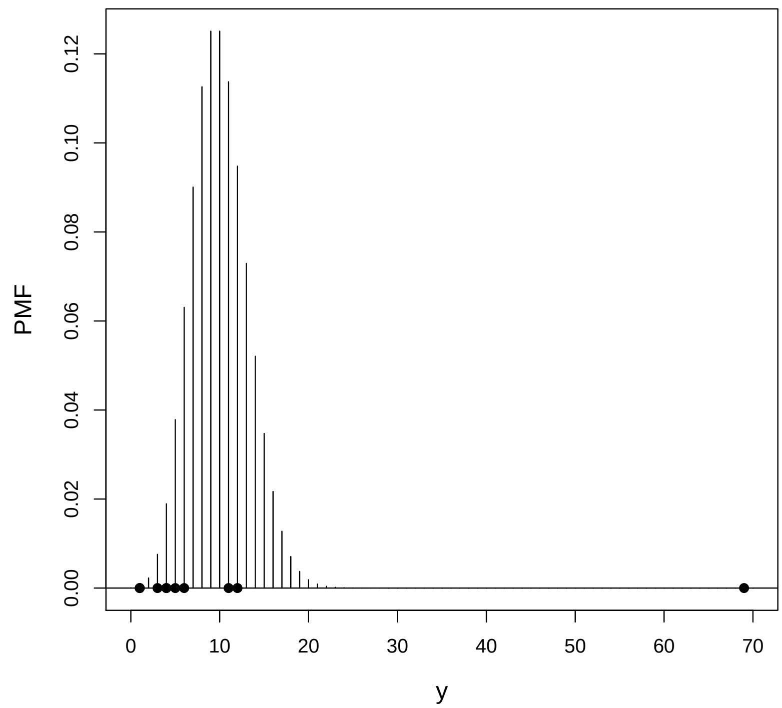
$$\frac{p_Y(y)}{p_Y(y+1)} > 1,$$

for all $y \in \{1, 2, 3, ...,\}$. When $y = 1$, we would have

$$\frac{p_Y(1)}{p_Y(2)} > 1 \implies p_Y(1) > p_Y(2);$$

when $y = 2$, we would have

$$\frac{p_Y(2)}{p_Y(3)} > 1 \implies p_Y(2) > p_Y(3);$$



and so on. That is, showing $p_Y(y)/p_Y(y+1) > 1$ for all $y \in \{1, 2, 3, ..., \}$ would imply

$$p_Y(1) > p_Y(2) > p_Y(3) > p_Y(4) > \cdots,$$

demonstrating that $y = 1$ is the most likely value. Suppose $y \in \{1, 2, 3, ..., \}$. We have

$$\frac{p_Y(y)}{p_Y(y+1)} = \frac{\dfrac{-(1-p)^y}{y \ln p}}{\dfrac{-(1-p)^{y+1}}{(y+1)\ln p}} = \frac{y+1}{y} \times \frac{(1-p)^y}{(1-p)^{y+1}} = \frac{y+1}{y} \times \frac{1}{1-p}.$$

We know that $y+1 > y$ so $(y+1)/y > 1$. Similarly, $0 < p < 1$ implies $1-p < 1$ so $1/(1-p) > 1$. The product of two numbers each larger than 1 is larger than 1. □

Remark: You could also regard $p_Y(y)$ to be a continuous function of y and then show $(d/dy)p_Y(y) < 0$ for all $y \geq 1$. This would show $p_Y(y)$ is a decreasing function of y and hence $p_Y(y)$ must be at its maximum value when $y = 1$.

(b) The first derivative of $m_Y(t)$ is

$$\frac{d}{dt} m_Y(t) = \frac{d}{dt}\left[\frac{\ln(1-qe^t)}{\ln p}\right] = \frac{1}{\ln p} \times \frac{-qe^t}{1-qe^t}.$$

Therefore,

$$E(Y) = \left.\frac{d}{dt} m_Y(t)\right|_{t=0} = \frac{1}{\ln p} \times \frac{-qe^0}{1-qe^0} = \frac{1}{\ln p} \times \frac{-q}{1-q} = -\frac{1-p}{p \ln p}.$$

The second derivative of $m_Y(t)$ is

$$\begin{aligned}\frac{d^2}{dt^2}m_Y(t) = \frac{d}{dt}\left(\frac{1}{\ln p}\times\frac{-qe^t}{1-qe^t}\right) &= -\frac{1}{\ln p}\times\frac{d}{dt}\left(\frac{qe^t}{1-qe^t}\right)\\ &= -\frac{1}{\ln p}\left[\frac{qe^t(1-qe^t)-qe^t(-qe^t)}{(1-qe^t)^2}\right]\\ &= -\frac{1}{\ln p}\left[\frac{qe^t-(qe^t)^2+(qe^t)^2}{(1-qe^t)^2}\right] = -\frac{1}{\ln p}\left[\frac{qe^t}{(1-qe^t)^2}\right].\end{aligned}$$

Therefore,

$$E(Y^2) = \frac{d^2}{dt^2}m_Y(t)\bigg|_{t=0} = -\frac{1}{\ln p}\left[\frac{qe^0}{(1-qe^0)^2}\right] = -\frac{1}{\ln p}\left[\frac{q}{(1-q)^2}\right] = -\frac{1-p}{p^2\ln p}.$$

(c) Showing

$$E(Y) = -\frac{1-p}{p\ln p}$$

without using the mgf is done by using the definition of mathematical expectation. We have

$$E(Y) = \sum_{y=1}^{\infty} yp_Y(y) = \sum_{y=1}^{\infty} y\times\frac{-(1-p)^y}{y\ln p} = -\frac{1}{\ln p}\sum_{y=1}^{\infty}(1-p)^y = -\frac{1}{\ln p}\left[\sum_{y=0}^{\infty}(1-p)^y - 1\right].$$

We recognize $\sum_{y=0}^{\infty}(1-p)^y$ as an infinite geometric sum with common ratio $r = 1-p$. Because $0 < 1-p < 1$, this sum converges and

$$\sum_{y=0}^{\infty}(1-p)^y = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Therefore,

$$E(Y) = -\frac{1}{\ln p}\left(\frac{1}{p}-1\right) = -\frac{1}{\ln p}\left(\frac{1-p}{p}\right) = -\frac{1-p}{p\ln p}.$$