Midterm 2 — Solutions

1. Assume everyone in Madison owns between zero and three cats and between zero and two dogs. Let X be the number of cats (Y be the number of dogs) a randomly chosen person owns. Suppose the joint pmf of X and Y is given by

		Cats			
		0	1	2	3
Dogs	0	12/48	6/48	3/48	3/48
	1	8/48	4/48	2/48	2/48
	2	4/48	2/48	1/48	1/48

(a) Find marginal distribution of X and Y.

$$X : \begin{array}{c|c|c|c} 0 & 1 & 2 & 3 \\ \hline 1/2 & 1/4 & 1/8 & 1/8 \end{array} \qquad Y : \begin{array}{c|c|c} 0 & 1 & 2 \\ \hline 1/2 & 1/3 & 1/6 \end{array}$$

(b) Are X and Y independent? Justify your answer mathematically.
Yes, because our joint distribution can be written as:

		Cats			
		0	1	2	3
Dogs	0	$1/2 \cdot 1/2$	$1/4 \cdot 1/2$	$1/8 \cdot 1/2$	$1/8 \cdot 1/2$
	1	$1/2 \cdot 1/3$	$1/4 \cdot 1/3$	$1/8 \cdot 1/3$	$1/8 \cdot 1/3$
	2	$1/2 \cdot 1/6$	$1/4 \cdot 1/6$	$1/8 \cdot 1/6$	$1/8 \cdot 1/6$

(c) Let $Z = X + Y$ be the totally number of pets (cats and dogs) the randomly chosen person owns. Compute $E(Z)$. One way to do it is to say that

$$E(X+Y) = E(X) + E(Y) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = \frac{37}{24}.$$

2. Let X be a random variable with moment generating function $M(t) = \left(1 - \frac{t}{2}\right)^{-2}$ when $|t| < 2$. What is the variance of X?

$$E(X) = M'(0) = \left(\frac{1}{(1-\frac{t}{2})^2}\right)'\Big|_{t=0} = \frac{1}{(1-\frac{t}{2})^3}\Big|_{t=0} = 1;$$

$$E(X^2) = M''(0) = \left(\frac{1}{(1-\frac{t}{2})^2}\right)''\Big|_{t=0} = \left(\frac{1}{(1-\frac{t}{2})^3}\right)'\Big|_{t=0} = \frac{3/2}{(1-\frac{t}{2})^4}\Big|_{t=0} = \frac{3}{2};$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{3}{2} - 1 = \frac{1}{2}.$$

3. The owner of a certain website is studying the distribution of the number of visitors to the site. Every day, a million people independently decide whether to visit the site, with probability $p = 2 \times 10^{-6}$ of visiting. Give a good approximation for the probability of getting at least three visitors on a particular day.

   Let $X \sim Bin(n, p)$ be the number of visitors, where $n = 10^6$. Since n is large, p is small, and $np = 2$ is moderate, $Pois(2)$ is a good approximation. This gives

$$P(X \geq 3) = 1 - P(X < 3) \approx 1 - e^{-2} - 2 \cdot e^{-2} - \frac{2^2}{2!}e^{-2} = 1 - 5e^{-2} \approx 0.3233.$$

4. Sides of fair dice are 2, 2, 3, 4, 5, 6. Roll the dice four times.

   (a) Find the probability that you roll 2 once, 3 once, and 4 twice (at some order). Simplify your answer.

$$\frac{4!}{2!1!1!}\left(\frac{2}{6}\right)^1\left(\frac{1}{6}\right)^1\left(\frac{1}{6}\right)^2 = \frac{24 \cdot 2}{2 \cdot 6^4} = \frac{4}{6^3} = \frac{1}{54}.$$

   (b) Find the probability that you roll 6 exactly once.

$$\binom{4}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^3.$$

   (c) Approximate the probability that, after 450 rolls, you have rolled 130 or more 2's.

   Let X be the number of "2". X binomial distribution $Bin(450, \frac{2}{6})$. We can approximate it by normal distribution ($np > 10$):

$$P(X > 130) = P\left(\frac{X - 150}{\sqrt{450 \cdot \frac{1}{3}\frac{2}{3}}} > \frac{130 - 150}{\sqrt{450 \cdot \frac{1}{3}\frac{2}{3}}}\right) == P\left(\frac{X - 150}{10} > -2\right) \approx 1 - \Phi(-2) = \Phi(2) \approx 0.9772$$

5. Let X and Y be random variables with joint density

$$f(x, y) = \begin{cases} cx & (x, y) \in [0, 1] \times [0, 2] \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Find the constant c.

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\, dxdy = \int_0^1\int_0^2 cx\, dxdy = c\int_0^1 x\, dx \int_0^2 dy = c\frac{1}{2} \cdot 2 = 1 \Leftrightarrow c = 1$$

(b) Find the marginal densities for X and Y.

We can do it by integration, but can say that

$$f(x,y) = \begin{cases} 2x \cdot \frac{1}{2} & (x,y) \in [0,1] \times [0,2] \\ 0 & \text{otherwise} \end{cases},$$

so, marginal densities should be

$$f_X(x) = \begin{cases} 2x & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}, \qquad f_Y(y) = \begin{cases} \frac{1}{2} & y \in [0,2] \\ 0 & \text{otherwise} \end{cases},$$

(and X and Y are independent).

(c) Find $P(X > Y)$.

Draw diagram!

$$P(X > Y) = \int_0^1 dx \int_0^x x dy = \int_0^1 x\, dx \int_0^x dy = \int_0^1 x^2\, dx = \frac{1}{3}.$$

6. Suppose that X has probability density function given by $\frac{1}{x^2}$ for $x > 1$ and 0 otherwise. Find the pdf of $|X-3|$.

Note that $Y = |X-3|$ will always be non-negative. We will use the cdf method to find the pdf of Y. Assume $y \geq 0$, then

$$F_Y(y) = P(Y \leq y) = P(|X-3| \leq y) = P(-y \leq X-3 \leq y) = P(3-y \leq X \leq 3+y) = F_X(3+y) - F_X(3-y).$$

Differentiating in y gives

$$f_Y(y) = F_Y'(y) = F_X'(3+y) + F_X'(3-y) = f_X(3+y) + f_X(3-y).$$

Since $3+y > 1$, we have $f_X(3+y) = 1/(3+y)^2$ for all $y \geq 0$. On the other hand, $f_X(3-y)$ will be nonzero only if $1 < 3-y$ which is equivalent to $y < 2$. In that case $f_X(3-y) = 1/(3-y)^2$. This gives

$$f_Y(y) = \begin{cases} \dfrac{1}{(3+y)^2} + \dfrac{1}{(3-y)^2}, & \text{if } 0 < y < 2 \\ \dfrac{1}{(3+y)^2}, & \text{if } 2 \leq y \\ 0 & \text{if } y \leq 0 \end{cases}$$

7. Assume a baseball team wins any given baseball game with probability p (regardless of who there opponent is). Suppose they played 100 games and won 60 of them. How confident are you that the true probability of winning a game is within 0.1 of the observed probability of winning a game $\hat{p} = 0.6$?

As we know

$$P(|\hat{p} - p| < 0.1) \geq 2\Phi(2 \cdot 0.1 \cdot \sqrt{100}) - 1 = 2\Phi(2) - 1 \approx 2 \cdot 0.9772 = 0.9544.$$