Math 5010 § 1. **Second Midterm Exam** Name: Solutions

Treibergs $\sigma\tau$ April 1, 2009

---

1. *Let X be a random variables satisfying the Poisson distribution with parameter $\lambda > 0$, and pmf $f_X(x) = e^{-\lambda}\,\lambda^x/x!$ for $x \in D = \{0, 1, 2, \ldots\}$. (a.)Find $\mathbf{E}(e^X)$. (b.)Let $Y = g(X)$ where $g(x) = 3$ if $x \le 2$ and $g(x) = 4$ if $x > 2$. Find the pmf $f_Y(y)$ for Y.*

   (a.) The expectation of a function of a random variable is $\mathbf{E}(g(X)) = \sum_x g(x)\, f_X(x)$. Thus

$$\mathbf{E}(e^X) = \sum_{x=0}^{\infty} e^x\, e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e\lambda)^x}{x!} = e^{-\lambda} e^{e\lambda} = e^{e\lambda - \lambda}.$$

   (b.) The pmf of a function of a random variable is given by $f_Y(y) = \sum_{x,\, g(x)=y} f_X(x)$ so

$$f_Y(3) = \sum_{x,\ g(x)=3} f_X(x) = \sum_{x=0}^{2} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right),$$
$$f_Y(4) = \sum_{x,\ g(x)=4} f_X(x) = \sum_{x=3}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!} = 1 - f_Y(3) = 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right)$$

   and $f_Y(y) = 0$ if $y \notin \{3, 4\}$.

2. *A standard die is rolled repeatedly. (a.)What is average number of rolls needed for the appearance of the first six? (b.)What is the probability that the second six occurs at the twelvth roll? (c.)What is the probability that the first six appears before the fifth roll?*

   (a.) If X is the number of rolls needed to get the first six, $X \sim \text{geom}(p)$ with $p = \frac{1}{6}$. Hence, $\mathbf{E}(X) = \frac{1}{p} = 6$.

   (b.) If Y is the number of rolls needed to first get two sixes, it has the negative binomial distribution with parameters $k = 2$ and $p = \frac{1}{6}$. Hence, $f_Y(y) = \binom{y-1}{k-1} p^k\, q^{y-k}$ so

$$\mathbf{P}(Y = 12) = f_Y(12) = \binom{11}{1} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{10} = \frac{11 \cdot 5^{10}}{6^{12}} \approx 0.0493.$$

   (c.) The probability that the first six happens before the fifth roll is $\mathbf{P}(X < 5) = \sum_{x=1}^{4} f_X(x)$ where $f_X(x) = pq^{x-1}$. Thus

$$\mathbf{P}(X < 5) = f_X(1) + f_X(2) + f_X(3) + f_X(4) = p + pq + pq^2 + pq^3 = \frac{p(1 - q^4)}{1 - q} = 1 - q^4.$$

   Alternatively, $\mathbf{P}(X < 5) = 1 - \mathbf{P}(X \ge 5) = 1 - \mathbf{P}(\text{no six in 4 rolls}) = 1 - q^4$. For $p = \frac{1}{6}$ this is $\mathbf{P}(X < 5) \approx 0.518$.

3. *An urn has three balls numbered 0, 1, 2. Let X be the number on a randomly chosen ball. Then flip X fair coins and let Y be the number of heads. The joint pmf is given in the table. Find $\mathbf{E}(X \mid Y)$.*

   By summing rows, we find the marginal pmf $f_Y(y) = \sum_x f(x,y)$. Thus

$f(x,y)$	$x=0$	$x=1$	$x=2$	$f_Y(y)$
$y=0$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{7}{12}$
$y=1$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{4}{12}$
$y=2$	0	0	$\frac{1}{12}$	$\frac{1}{12}$

   The conditional expectation is given by $\mathbf{E}(X \mid Y) = \sum_x x\, f_{X \mid Y}(x \mid Y)$ which is a random variable (function of y). Using the formula for conditional pmf,

$$f_{X \mid Y}(x|y) = \mathbf{P}(X = x|Y = y) = \frac{\mathbf{P}(X = x \text{ and } Y = y)}{\mathbf{P}(Y = y)} = \frac{f(x,y)}{f_Y(y)}.$$

   Thus, substituting table values,

$$\begin{aligned}
\mathbf{E}(X \mid Y = 0) &= 0 \cdot \frac{f(0,0)}{f_Y(0)} + 1 \cdot \frac{f(1,0)}{f_Y(0)} + 2 \cdot \frac{f(2,0)}{f_Y(0)} = \frac{1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{12}}{\frac{7}{12}} = \frac{4}{7};\\
\mathbf{E}(X \mid Y = 1) &= 0 \cdot \frac{f(0,1)}{f_Y(1)} + 1 \cdot \frac{f(1,1)}{f_Y(1)} + 2 \cdot \frac{f(2,1)}{f_Y(1)} = \frac{1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6}}{\frac{4}{12}} = \frac{3}{2};\\
\mathbf{E}(X \mid Y = 2) &= 0 \cdot \frac{f(0,2)}{f_Y(2)} + 1 \cdot \frac{f(1,2)}{f_Y(2)} + 2 \cdot \frac{f(2,2)}{f_Y(2)} = \frac{1 \cdot 0 + 2 \cdot \frac{1}{12}}{\frac{1}{12}} = 2.
\end{aligned}$$

4. *Suppose k distinguishable balls are distributed to m distinguishable urns in such a way that each ball is equally likely to go into any urn (and an urn may contain more than one ball). What is the expected number of occupied urns? Notice that $X = X_1 + X_2 + \cdots + X_m$ with X_i being the random variable equal to 1 if the i^{th} urn is occupied and 0 otherwise.*

   Assume that the urn chosen for each ball is independent. Then the probability that any given ball is not in the ith urn is $(m-1)/m$ and

$$\mathbf{E}(X_i) = \mathbf{P}(i\text{th urn is occupied}) = 1 - \mathbf{P}(\text{no ball in } i\text{th urn}) = 1 - \left(\frac{m-1}{m}\right)^k.$$

   Hence

$$\mathbf{E}(X) = \sum_{i=1}^{m} \mathbf{E}(X_i) = m\left[1 - \left(\frac{m-1}{m}\right)^k\right].$$

5. *Let $X_1, X_2, X_3, \ldots$ be a sequence of mutually independent random variables. Suppose that each X_n is uniformly distributed on $\{1, 2, 3, \ldots, k\}$. How big does the random sample have to be so that you are 90% sure that your sample mean lies between $\frac{k}{2}$ and $1 + \frac{k}{2}$? In other words, how big should n be so that $\mathbf{P}\left(\frac{k}{2} < \frac{1}{n}(X_1 + X_2 + \cdots + X_n) < 1 + \frac{k}{2}\right) \geq 0.9$?*

   Observe that a random variable X_n that is uniformly distributed on $\{1, \ldots, k\}$ has

$$\mathbf{E}(X_n) = \frac{k+1}{2}, \qquad \mathbf{Var}(X_n) = \frac{k^2-1}{12}.$$

Put

$$S_n = X_1 + X_2 + \cdots + X_n.$$

Now, because of independence,

$$\mathbf{E}\left(\frac{1}{n}S_n\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(X_i) = \frac{k+1}{2}, \qquad \mathbf{Var}\left(\frac{1}{n}S_n\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{Var}(X_i) = \frac{k^2-1}{12n}. \tag{1}$$

Chebychov's Inequality for the random variable Z and $a > 0$ is

$$\mathbf{P}\left(|Z - \mathbf{E}(Z)| \geq a\right) \leq \frac{\mathbf{Var}(Z)}{a^2}.$$

Applying Chebychov's inequality to $Z = \frac{1}{n}S_n$ and $a = \frac{1}{2}$ and using (1) we find

$$\begin{aligned}
\mathbf{P}\left(\frac{k}{2} < \frac{X_1 + X_2 + \cdots + X_n}{n} < 1 + \frac{k}{2}\right) &= \mathbf{P}\left(-\frac{1}{2} < \frac{1}{n}S_n - \frac{k+1}{2} < \frac{1}{2}\right) \\
&= \mathbf{P}\left(\left|\frac{1}{n}S_n - \frac{k+1}{2}\right| < \frac{1}{2}\right) \\
&= 1 - \mathbf{P}\left(\left|\frac{1}{n}S_n - \frac{k+1}{2}\right| \geq \frac{1}{2}\right) \\
&\geq 1 - \frac{k^2-1}{12n\left(\frac{1}{2}\right)^2} = 1 - \frac{k^2-1}{3n}
\end{aligned}$$

We need n big enough so that

$$1 - \frac{k^2-1}{3n} \geq 0.9 = 1 - 0.1$$

or

$$n \geq \frac{10}{3}(k^2-1).$$