Math 477: Mathematical Theory Of Probability

Second Exam Solution

August 4th, 2006

Name: Eduardo Osorio

1. **(15 pts)** In a town of 50,000 people, each person buys one lottery ticket. The chances of any particular ticket's being a winner are 1 in 100,000. Give an approximation for the probability that no one in the town has a winning ticket, using an appropiate approximation scheme (state clearly your assumptions and what your scheme is!). Write your answer in decimal form. Compare to the "exact" probability.

   Sln: Let X = Number of winners in the town. Let's assume that each person buys a ticket independently from the rest. In this case, $X \sim \text{Bin}\big(50000, \frac{1}{100000}\big)$.
   We want an approximation for $\mathbb{P}(X = 0)$. Since $np(1-p) = 0.499995 < 10$, it is better to use a Poisson random variable to approximate this probability. Let $Y \sim \text{Poisson}(0.5)$. Hence,

   $$\mathbb{P}(X = 0) \simeq \mathbb{P}(Y = 0) = e^{-.5} \simeq 0.60653066$$

   To get the exact probability, we recall that

   $$\mathbb{P}(X = 0) = \left(1 - \frac{1}{100000}\right)^{50000} \simeq 0.60652914$$

   and we confirm that in fact this is a pretty good approximation.

2. A discrete random variable X has probability mass function

$$p_X(x) = \begin{cases} 0.2 & \text{if } x = -3 \\ 0.3 & \text{if } x = -2 \\ 0.1 & \text{if } x = -1 \\ c & \text{if } x = 1 \\ 0.25 & \text{if } x = 3 \\ 0 & \text{if } x \neq -3, -2, -1, 1, 3 \end{cases}$$

(a) **(5 pts)** Find the value of c.

Sln: From the definition of $p_X(\cdot)$ we see that X takes values atmost on the set $\{-3, -2, -1, 1, 3\}$, probabilistically. Therefore, since the sum of the mass function at all the possible values has to be one,

$$c = 1 - 0.2 - 0.3 - 0.1 - 0.25 = 0.15$$

(b) **(5 pts)** Find $\mathbb{E}[X]$.

Sln: By definition,

$$\mathbb{E}[X] = (-3)(0.2)+(-2)(0.3)+(-1)(0.1)+1(0.15)+3(0.25) = -0.4$$

(c) **(5 pts)** Find $\text{Var}(X)$.

Sln: We know that $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. It is enough then to calculate $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = 9(0.2) + 4(0.3) + 1(0.1) + 1(0.15) + 9(0.25) = 5.5$$

Thus, $\text{Var}(X) = 5.5 - (-0.4)^2 = 5.34$

3. A fair die is tossed 6000 times.

(a) **(5 pts)** Let X be the number of times a 5 or a 6 appears on top. Name the distribution of the random variable X with its correspondent parameter(s).

Sln: X is a binomial random variable with parameters $n = 6000$ and $p = \frac{2}{6} = \frac{1}{3}$. That is, $X \sim \text{Bin}\big(6000, \frac{1}{3}\big)$.

(b) **(5 pts)** Calculate the exact probability that a 5 or a 6 appears on top more than 2006 times. Give your answer in combinatorial notation.

Sln: We want to calculate $\mathbb{P}(X > 2006)$. That is,

$$\mathbb{P}(X > 2006) = \sum_{i=2007}^{6000} p_X(i) = \sum_{i=2007}^{6000} \binom{6000}{i} \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{6000-i}$$

(c) **(15 pts)** Approximate the probability that a 5 or a 6 appears on top more than 2006 times. Give your answer in decimal form.

Sln: Since $np(1-p) = 6000\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{12000}{9} > 10$ ($np = 2000$), a Normal approximation is appropriate (don't forget the correction for continuity):

$$\begin{aligned}
\mathbb{P}(X > 2006) &= \mathbb{P}(X \geq 2007) \\
&= \mathbb{P}(X \geq 2006.5) \\
&= \mathbb{P}\left(\frac{X - 2000}{\sqrt{\frac{12000}{9}}} \geq \frac{2006.5 - 2000}{\sqrt{\frac{12000}{9}}}\right) \\
&\simeq \mathbb{P}\left(Z \geq 0.178\right) \\
&\simeq 1 - \Phi(0.18) = 0.4286
\end{aligned}$$

where $Z \sim \text{Normal}(0, 1)$.

Note: I took 3 points off if you didn't justify why the appropiate way to approximate this probability was by using a Normal random variable. I'll add 3 points to **everybody** in the exam grade, but I advise you to answer the questions justifying everything you do... not only for what it remains of the course but also for your future exams... especially when who is grading grades the procedure and not only your final answer.

4. Suppose that the density function of a continuous random variable X is

$$f_X(x) = \begin{cases} \frac{1}{8}(3x+4) & \text{if } -1 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) **(10 pts)** Find the cumulative distribution function F_X. Be sure to give the values of $F_X(x)$ for all real x.

Sln: By definition,

$$F_X(a) = \mathbb{P}(X \le a) = \int_{-\infty}^{a} f_X(x)dx$$

If $a < -1$, $f_X(x)$ will be zero for all x before a, thus

$$F_X(a) = \int_{-\infty}^{a} 0dx = 0$$

If $-1 \le a < 1$, since $f_X(x) = 0$ for $x < -1$, we get that

$$F_X(a) = \int_{-\infty}^{a} f_X(x)dx = \int_{-1}^{a} \frac{1}{8}(3x+4)dx = \frac{1}{8}\left(\frac{3}{2}x^2 + 4x\right)_{-1}^{a} = \frac{1}{8}\left(\frac{3}{2}a^2 + 4a + \frac{5}{2}\right)$$

If $a \ge 1$,

$$F_X(a) = \int_{-\infty}^{a} f_X(x)dx = \int_{-1}^{1} \frac{1}{8}(3x+4)dx = 1$$

(b) **(10 pts)** Find $\mathbb{P}\left(\left|X - \frac{1}{2}\right| > 1\right)$.

Sln: Note that $|X - 1/2| > 1$ iff $X > 3/2$ or $X < -1/2$. Thus,

$\mathbb{P}(|X-1/2| > 1) = \mathbb{P}(X > 3/2) + \mathbb{P}(X < -1/2) = 0 + F_X(-1/2) = 7/64$

(c) **(5 pts)** Find $\mathbb{E}[X]$.

Sln: By definition,

$$\mathbb{E}[X] = \int_{-1}^{1} x\frac{1}{8}(3x+4)dx = 1/4$$

(d) **(5 pts)** Find Var(X).

Sln: We know that $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. Let's get $\mathbb{E}[X^2]$ first...

$$\mathbb{E}[X^2] = \int_{-1}^{1} x^2 \frac{1}{8}(3x+4)dx = 1/3$$

Thus, $\mathrm{Var}(X) = 1/3 - 1/16 = 13/48$

5. The experts at NJT (New Jersey Transit) say that the average number of accidents occurring on the turnpike each day is 3.

   Explain your reasoning!

   (a) **(10 pts)** What is the probability that three or more accidents occur today on the turnpike?

   Sln: We can idealize the accidents on the turnpike as events happening along a time line that follow the three assumptions we had for defining a Poisson process (see page 170, Ross' book). In this case, let X be the number of accidents that occur on the turnpike in a day... we know that $X \sim \text{Poisson}(\lambda)$ for some $\lambda > 0$, but we are told that $\mathbb{E}[X] = 3$. Thus $X \sim \text{Poisson}(3)$.
   We want to calculate $\mathbb{P}(X \geq 3)$. Hence,

$$\begin{aligned}
\mathbb{P}(X \geq 3) &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2) \\
&= 1 - e^{-3}(1 + 3 + 3^2/2) \\
&= 1 - \frac{13}{2}e^{-3}
\end{aligned}$$

   (b) **(5 pts)** What is the probability that three or more accidents occur today under the assumption that at least one accident occurs today?

   Sln: By definition of conditional probability,

$$\begin{aligned}
\mathbb{P}(X \geq 3 | X \geq 1) &= \frac{\mathbb{P}(X \geq 3, X \geq 1)}{\mathbb{P}(X \geq 1)} \\
&= \frac{\mathbb{P}(X \geq 3)}{\mathbb{P}(X \geq 1)} \\
&= \frac{1 - \frac{13}{2}e^{-3}}{1 - e^{-3}}
\end{aligned}$$

6. **(Extra problem, 10 pts)** Find a collection of events $\{E_a\}$, $0 < a < 1$, having the property that $\mathbb{P}(E_a) = 1$ for all a, but

$$\mathbb{P}\left(\bigcap_a E_a\right) = 0 \quad [!]$$

(Does this contradict a result from another extra problem I gave you before?)
Hint: Let X be uniform over $(0, 1)$ and define each E_a in terms of X.

Sln: Let $X \sim \text{Uniform}((0, 1))$ and define $E_a := \{X \neq a\}$ for each a in the interval $(0, 1)$. Since X is a continuous random variable we know that $\mathbb{P}(X \neq a) = 1$ for all a. Also, when we take the intersection of **all** the E_a's we get the empty set, why?. Therefore,

$$\mathbb{P}\left(\bigcap_a E_a\right) = 0$$

Note this doesn't contradict an extra problem I gave you before because the collection of sets here is **not** countable.

Eduardo Osorio