Introduction to Probability, Fall 2013

Math 30530 Section 01

Homework 10 — solutions

1. (a) Let X be a uniformly selected random number on the interval $[0, 1]$. For $a > 0$ and $b \in \mathbb{R}$, let $Y = aX + b$. Calculate the density function of Y.

   Solution: Possible values for Y: anything from b to $a+b$. For each $b \leq y \leq a+b$,

$$\Pr(Y \leq y) = \Pr(aX + b \leq y) = \Pr(X \leq (y-b)/a) = (y-b)/a,$$

   the last equality using the fact that X is uniform on $[0, 1]$. So, the CDF of Y is

$$F_Y(y) = \begin{cases} 0 & \text{if } y < a, \\ (y-b)/a & \text{if } a \leq y \leq b+a, \\ 1 & \text{if } y > b+a, \end{cases}$$

   and the density of Y is

$$f_Y(y) = \begin{cases} 0 & \text{if } y < a, \\ 1/a & \text{if } a \leq y \leq b+a, \\ 0 & \text{if } y > b+a. \end{cases}$$

   (b) Write down the density function of a uniformly selected random number on the interval $[b, a+b]$ ($a > 0$ and $b \in \mathbb{R}$).

   Solution: Exactly the same as the density of Y in the last part: if X is a uniformly selected random number on the interval $[b, a+b]$, then the density of X is

$$f_X(x) = \begin{cases} 0 & \text{if } x < a, \\ 1/a & \text{if } a \leq x \leq b+a, \\ 0 & \text{if } x > b+a. \end{cases}$$

2. I throw a dart n times at a dartboard with radius 1, each time selecting a uniform and independent point from the board. Let X_i be the random variable that records the distance from my ith throw to the center of the dartboard, and let $Y_{(n)}$ be the distance to the center of the dartboard of my *closest* throw (i.e.

$$Y_{(n)} = \min\{X_1, \ldots, X_n\}).$$

(a) Find the density function of $Y_{(n)}$.

Solution: We start with the CDF of $Y_{(n)}$. The possible values of $Y_{(n)}$ are anything from 0 to 1, so we each $0 \le y \le 1$ we want to compute $\Pr(Y_{(n)} \le y)$. It's easier to compute $\Pr(Y_{(n)} \ge y)$, because

$$\begin{aligned}
\Pr(Y_{(n)} \ge y) &= \Pr(\min\{X_1, \ldots, X_n\} \ge y) \\
&= \Pr(X_1 \ge y \text{ AND } X_2 \ge y \text{ AND } \ldots \text{ AND } X_n \ge y) \\
&= \left(\Pr(X_1 \ge y)\right)^n .
\end{aligned}$$

The probability that $X_1 \ge y$ is the probability of landing in the annulus outside the circle of radius y, which is $(\pi 1^2 - \pi y^2)/(\pi 1^2) = 1 - y^2$. So

$$\Pr(Y_{(n)} \le y) = 1 - \Pr(Y_{(n)} \ge y) = 1 - (1 - y^2)^n.$$

It follows that the CDF of $Y_{(n)}$ is

$$F_n(y) = \begin{cases} 0 & \text{if } y < 0, \\ 1 - (1 - y^2)^n & \text{if } 0 \le y \le 1, \\ 1 & \text{if } y > 1, \end{cases}$$

and the density of $Y_{(n)}$ is

$$f_n(y) = \begin{cases} 0 & \text{if } y < 0, \\ 2yn(1 - y^2)^{n-1} & \text{if } 0 \le y \le 1, \\ 0 & \text{if } y > 1. \end{cases}$$

(b) For $n = 1, 2, 3, 4$, find $E(Y_{(n)})$.

Solution: From the density, $E(Y_{(n)}) = \int_0^1 2y^2 n(1 - y^2)^{n-1}\, dy$. Calculating this integral for $n = 1, 2, 3, 4$ gives values of 2/3, 8/15, 16/35, and 128/315 (roughly .67, .53, .46, and .41).

(c) For $n = 1, 2, 3, 4$, find $\Pr(Y_{(n)} < .5)$.

Solution: From the density, $\Pr(Y_{(n)} < .5) = \int_0^{.5} 2yn(1 - y^2)^{n-1}\, dy$. Calculating this integral for $n = 1, 2, 3, 4$ gives values of .25, .4375, .578125, and .68359375.

3. Use the transform of the exponential random variable (which we calculated in class) to compute $E(X)$ and $\mathrm{Var}(X)$ when $X \sim \text{exponential}(\lambda)$.

Solution: We computed $M_X(s) = \frac{\lambda}{\lambda - s}$ for $X \sim \text{exponential}(\lambda)$ (as long as $s < \lambda$), so

$$M_x'(s) = \frac{\lambda}{(\lambda - s)^2}, \quad \text{so} \quad E(X) = M_x'(0) = \frac{1}{\lambda},$$

and

$$M_x''(s) = \frac{2\lambda}{(\lambda - s)^3}, \quad \text{so} \quad E(X^2) = M_x''(0) = \frac{2}{\lambda^2},$$

so

$$\mathrm{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

4. (a) Let $X \sim \text{Poisson}(\lambda)$. Calculate the transform X.

Solution: $P(K = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ for $X \sim \text{Poisson}(\lambda)$ $(k = 0, 1, 2, 3, \ldots)$, so

$$\begin{aligned} M_X(s) &= E(e^{sX}) \\ &= \sum_{k=0}^{\infty} e^{sk}\frac{\lambda^k}{k!}e^{-\lambda} \\ &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^s\lambda)^k}{k!} \\ &= e^{-\lambda}e^{e^s\lambda} \\ &= e^{\lambda(e^s-1)}, \end{aligned}$$

the second-from-last inequality using the power series for e^x, $e^x = 1 + x + \ldots + x^l/k! + \ldots$. This transform is valid for all s.

(b) Let $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$ be independent. Use transforms to show that $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Solution: From previous part,

$$M_X(s) = e^{\lambda_1(e^s-1)}, \qquad M_Y(s) = e^{\lambda_2(e^s-1)}.$$

Since we know that the transform of a sum of independent rvs is a product of the transforms, we get

$$M_{X+Y}(s) = e^{\lambda_1(e^s-1)}e^{\lambda_2(e^s-1)} = e^{\lambda_1(e^s-1)+\lambda_2(e^s-1)} = e^{(\lambda_1+\lambda_2)(e^s-1)}.$$

This is exactly the transform of a Poisson random variable with parameter $\lambda_1+\lambda_2$, and so we are done.

5. I choose r items from a collection of $N + M$ items, one after the other, *without replacement*. N of the items are "good" and the remaining M are "bad". Let X_i be the indicator random variable indicating whether the ith item I chose was good (so $X_i = 1$ if the ith item was good, and $X_i = 0$ if it was bad). For $i \neq j$, calculate the covariance $\text{Cov}(X_i, X_j)$, and the correlation coefficient. (**Note**: it should be very small, going to 0 as N and M go to infinity; this justifies treating samples without replacement as being essentially independent when the population is large).

Solution: There are $N times (N + M - 1)(N + M - 2)\ldots(N + M - (r - 1))$ ways of choosing the r items so that the ith is good (think about choosing the ith item first, from among the N good items, and choosing the remaining $r-1$ items arbitrarily from what's left). There are $(N + M)(N + M - 1)(N + M - 2)\ldots(N + M - (r - 1))$ ways of choosing the r items, in total. So

$$\Pr(X_i = 1) = \frac{N \times (N + M - 1)(N + M - 2)\ldots(N + M - (r - 1))}{(N + M)(N + M - 1)(N + M - 2)\ldots(N + M - (r - 1))} = \frac{N}{N + M},$$

and $\Pr(X_i = 0) = M/(N + M)$. Similarly,

$$\Pr(X_j = 1) = \frac{N}{N + M},$$

and $\Pr(X_j = 0) = M/(N+M)$. We thus get that

$$E(X_i) = \frac{N}{N+M}, \quad E(X_i^2) = \frac{N}{N+M}, \quad \text{Var}(X_i) = \frac{MN}{(N+M)^2},$$

and

$$E(X_j) = \frac{N}{N+M}, \quad E(X_j^2) = \frac{N}{N+M}, \quad \text{Var}(X_j) = \frac{MN}{(N+M)^2}.$$

There are $N(N-1)times(N+M-2)(N+M-3)\ldots(N+M-(r-1))$ ways of choosing the r items so that the ith and jth are both good (think about choosing the ith item first, then the jth item, from among the N good items, and choosing the remaining $r-2$ items arbitrarily from what's left). There are $(N+M)(N+M-1)(N+M-2)\ldots(N+M-(r-1))$ ways of choosing the r items, in total. So

$$\begin{aligned}
\Pr(X_iX_j = 1) &= E(X_iX_j) \\
&= \frac{N(N-1) \times (N+M-2)\ldots(N+M-(r-1))}{(N+M)(N+M-1)(N+M-2)\ldots(N+M-(r-1))} \\
&= \frac{N(N-1)}{(N+M)(N+M-1)},
\end{aligned}$$

and so

$$\begin{aligned}
\text{Cov}(X_i, X_j) &= E(X_iX_j) - E(X_i)E(X_j) \\
&= \frac{N(N-1)}{(N+M)(N+M-1)} - \left(\frac{N}{N+M}\right)^2 \\
&= \frac{-NM}{(N+M)^2(N+M-1)}
\end{aligned}$$

and

$$\begin{aligned}
\rho &= \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)\text{Var}(X_j)}} \\
&= \frac{\frac{-NM}{(N+M)^2(N+M-1)}}{\frac{MN}{(N+M)^2}} \\
&= \frac{-1}{N+M-1}
\end{aligned}$$

This is indeed small; this shows that although X_i and X_j are very slightly (negatively) correlated, when $N+M$ is large they are essentially uncorrelated.

6. Chapter 4, problems 29 and 30 — see the supplementary solution file 1.

7. Chapter 4, problems 17, 18 and 19 — see the supplementary solution file 2.