MAT 3701: Selected Solutions to Assignments 5 & 6

Assignment 5: Additional Problems

1. "Let X and Y be discrete random variables, and let g and h be functions such that the following identity holds: $P(X = x \text{ and } Y = y) = g(x)h(y)$."

   Please note that, although the values of the joint density $f_{X,Y}(x, y)$ break down as products $g(x)h(y)$, g and h need *not* be the marginal densities of x and y. For example, if one were to multiply g by some factor and divide h by the same factor, the equation would still hold. There is no reason to believe that $\sum_x g(x) = 1$ or $\sum_y h(y) = 1$, and thus no reason to believe that g and h are probability densities at all (of any random variable).

   (a) "Express $P(X = x)$ in terms of g and h." $P(X = x) = \sum_y P(X = x \text{ and } Y = y) = \sum_y g(x)h(y) = g(x)\sum_y h(y)$.

   (b) "Express $P(Y = y)$ in terms of g and h." Similarly, $P(Y = y) = h(y)\sum_x g(x)$.

   (c) "Show that $(\sum_x g(x))(\sum_y h(y)) = 1$." Since $\sum_x P(X = x) = 1$, we have $1 = \sum_x g(x)\sum_y h(y) = (\sum_x g(x))(\sum_y h(y))$. (A similar calculation could have been done using the result for $P(Y = y)$ or, alternatively, by going back to the original joint density and using the fact that $\sum_{x,y} P((x, y)) = 1$.)

   (d) "Show that X and Y are independent." Use (a), (b), (c), and the definition of independent! $P(X = x, Y = y) = g(x)h(y) = g(x)(\sum_x g(x))(\sum_y h(y))h(y)$ [by (c)] $= \cdots$.

   [Recall that X and Y are independent if the joint density of X and Y is equal to the product of the densities of X and Y: $f(x, y) = f(x)f(y)$.]

(Source: *Introduction to Probability Theory* by Hoel, Port, and Stone, 1971, p. 79.)

Assignment 6: Additional Problems

1. As in a previous problem, a target is made by drawing concentric circles of radius 1, 2, 3, and 4 meters. Five darts are thrown at the target "at random"; that is, the darts are guaranteed to hit the target, and the probability a dart will land in any region is proportional to its area. Each throw is independent of the others.

   The "bull's eye" within the circle of radius 1 is assigned 4 points. The annular region between the circle of radius 1 and the circle of radius 2 is assigned 3 points. The annular region between the circle of radius 2 and the circle of radius 3 is assigned 2 points. The outermost annular region is assigned 1 point.

   (a) Let Y_i be the score obtained form the i^{th} dart, $i = 1, 2, 3, 4, 5$. Compute the density of Y_i. The total area of the target is 16π square meters, the area of the next concentric circle in is 9π square meters, etc. Thus, for each dart, the probability of hitting the bull's eye is $\frac{1}{16}$, the probability of hitting the next region is $\frac{4-1}{16} = \frac{3}{16}$, etc.

   (b) Compute $\mathrm{E}Y_i$. $4(\frac{1}{16}) + 3(\frac{3}{16}) + 2(\frac{5}{16}) + 1(\frac{7}{16}) = \frac{30}{16} = \frac{15}{8}$.

   (c) Let $\overrightarrow{Y} = (Y_1, Y_2, Y_3, Y_4, Y_5)$. Compute the density of $\overrightarrow{Y}$. Since the throws are independent, the probability of any sequence of throws is the product of their probabilities. (This is an acceptable answer; you don't need to write down all the formulas as long as you know how to find them.)

   (d) Let $X = \sum_{i=1}^{5} Y_i$. That is, X represents the total score from the five darts. Compute $\mathrm{E}X$. (Hint: There is an easy way!) The expectation of a sum is the sum of the expectations! So $\mathrm{E}X = \frac{75}{8}$.

2. We proved in class for any random variables X and Y and any real constant c that $\mathrm{E}(X + Y) = \mathrm{E}X + \mathrm{E}Y$ and that $\mathrm{E}cX = c\mathrm{E}X$.

   (a) Combine these results to conclude that, for any linear combination $aX + bY$ of random variables (where a and b are real constants), $\mathrm{E}(aX + bY) = a\mathrm{E} + b\mathrm{E}Y$. This should cause you no difficulty!

   (b) More generally, prove by induction that for any random variables $X_1, X_2, X_3, \ldots, X_n$ and real constants $c_1, c_2, c_3, \ldots, c_n$, $\mathrm{E}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i \mathrm{E}X_i$.

Proof. Initial claim (Case n=1): Both the case $n = 1$ and the case $n = 2$ have previously been proven directly.

Inductive claim: Assume as inductive hypothesis that

$$\mathrm{E}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i \mathrm{E} X_i.$$

Then $\mathrm{E}\left(\sum_{i=1}^{n+1} c_i X_i\right) = \mathrm{E}\left(\sum_{i=1}^{n} c_i X_i + c_{n+1} X_{n+1}\right)$, by the recursive definition of a sum with arbitrarily many terms. By the case of two terms (previously proven), we obtain that $\mathrm{E}\left(\sum_{i=1}^{n+1} c_i X_i\right) = \mathrm{E}\left(\sum_{i=1}^{n} c_i X_i\right) + \mathrm{E}(c_{n+1} X_{n+1})$. Finally, by inductive hypothesis, again using the recursive definition of a sum (in the opposite direction), we obtain the desired conclusion inductive that

$$\mathrm{E}\left(\sum_{i=1}^{n+1} c_i X_i\right) = \sum_{i=1}^{n+1} c_i \mathrm{E} X_i.$$

□

3. Suppose X, Y, and Z are random variables with finite expectation, and $EX = 2$, $EY = 7$, and $EZ = 3$. What is $E(2X + Y + 4Z)$? Explain.

$E(2X + Y + 4Z) = 2EX + EY + 4EZ = 23$. This calculation is a direct application of the result of the previous exercise.

4. "Let X be a geometrically distributed random variable [as defined by Hoel, Port, and Stone, slightly different from our text: $f(x) = (1 - p)^x p$] and let $M > 0$ be an integer. Set $Z = \min(X, M)$. Compute the mean [expectation] of Z." (Source: *Introduction to Probability Theory* by Hoel, Port, and Stone, 1971, p. 104.)

Remark: The interpretation of this version of the geometric distribution is that X tells you the number of failures that occur before the first success. If Y is a random variable with geometric distribution *as we have defined it*, then $X = Y - 1$, as you can see by substitution into the formula. I forgot to mention it, but implicit in this interpretation and distribution is that the possible values of X include 0.

This computation is not only good practice with the definition of expectation, but also with computing the sums of finite and infinite series! Since Z takes the minimum of the two values, $Z = X$ in the case that $X \le M$, and $Z = M$ in the case that $M < X$. Thus, using our general result on the expectation of a function of a random vector (here, Z is a function of X),

$$EZ = \sum_{0}^{M} x(1-p)^x p + \sum_{M+1}^{\infty} M(1-p)^x p.$$

The first of these two sums is probably the harder one, but we know from experience that our computation must be based on the following function: let $g(y) = \sum_0^M y^x$. Then $g'(y) = \sum_0^M x y^{x-1}$. We can find closed formulas for both of these sums, and of course we will evaluate $f'(1-p)$. So here goes! Using the usual trick for geometric sums, or just remembering the formula, we obtain $g(y) = \frac{1-y^{M+1}}{1-y}$. Computing $g'(y)$ from this formula is not pretty, but you all know how to do it. After some algebraic simplification, we obtain $g'(y) = \frac{My^{M+1} - (M+1)y^M + 1}{(1-y)^2}$; hence (again after some careful algebraic simplification, and check my work, because I don't guarantee there are no algebra errors!) $g'(1-p) = \frac{(1-p)^M(-Mp-1)+1}{p^2}$. Now, equipped with this systematic prior preparation, we obtain

$$\sum_{0}^{M} x(1-p)^x p = p(1-p)\sum_{0}^{M} x(1-p)^{x-1}$$

$$= p(1-p)g'(1-p) = \frac{(1-p)^{M+1}(-Mp-1) + 1 - p}{p}.$$

For the second sum, substituting $y = x - M - 1$ and using geometric series techniques, we obtain

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

I leave it to you to add these two sums together and simplify as appropriate.

5. "Let X be a geometrically distributed random variable [as defined by Hoel, Port, and Stone: $f(x) = (1-p)^x p$] and let $M > 0$ be an integer. Set $Z = \max(X, M)$. Compute the mean of Z." (Source: *Introduction to Probability Theory* by Hoel, Port, and Stone, 1971, p. 104.) As this computation uses similar methods to the one above, I leave it to you for practice if you did not solve it the first time.

6. Let X have hypergeometric density. Compute EX. Hint: The easiest way to do this is as follows, and is very similar to the method we used to compute the expectation of a binomial random variable. Let $X_i = 1$ if the i^{th} object picked is marked, and 0 if it is not marked. Thus, each X_i is Bernoulli. What is $P(X_i = 1)$? We then have $X = \sum_{i=1}^{n} X_i$, just as for a binomial random variable, except that the X_i are not independent in this case. Discussed in class.

7. Let X be a discrete random variable with probability distribution f. Let $\mu = \mathrm{E}X$, and let $Y = X - \mu$. Prove the E$Y = 0$. (Hint: Remember that $\sum_x f(x) = 1$.) There are two easy ways to do this:

   Method 1: use the result on sums of random variables. $EY = EX - \mu = 0$. Method 2: direct computation. $EY = \sum_x (x-\mu) f(x) = \sum_x x f(x) - \mu \sum_x f(x) = EX - \mu(1) = 0$. Incidently, in order to rearrange this series we have used the fact that if two series converge absolutely, the series of the sum of their terms converges absolutely. This is an elementary fact but not a trivial one; its proof requires some analysis.

8. Let X be a continuous random variable with probability density f. Let $\mu = \mathrm{E}X$, and let $Y = X - \mu$. Prove the E$Y = 0$. Similar to the previous exercise.

9. Let X be a continuous random variable with probability density f. Prove that if f is an even function, then E$X = 0$. (Recall that by definition a function f is *even* if, for all $x \in \mathbb{R}$, $f(-x) = f(x)$, and a function f is *odd* if, for all $x \in \mathbb{R}$, $f(-x) = -f(x)$. Hint: If f is an even function, what type of function is the one whose output at x is $xf(x)$?)

   Remark: We implicitly assume that the expectation of X is finite; otherwise the integrals in the computation below might not converge.

   If f is an even function, then $(-x)f(-x) = -xf(x)$, so the function with outputs $xf(x)$ is an odd function. Thus $\int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{0} xf(x)dx + \int_{0}^{\infty} xf(x)dx = -\int_{0}^{\infty} uf(u)du + \int_{0}^{\infty} xf(x)dx = 0$. (Make the substitution $u = -x$ in the first integral.)

10. Does the converse to the previous problem hold? If so, prove it; if not, give a counterexample. The converse does not hold. First, here is a simple discrete example, which I told the students who came to office hours. Let X take the value -2 with probability $\frac{1}{3}$ and the value 1 with probability $\frac{2}{3}$. Obviously, the probability distribution of X is not an even function (since, for example, $P(X = -1) = 0$, whereas $P(X = 1) = \frac{2}{3}$. Nonetheless, $EX = (-2)(\frac{1}{3} + (1)(\frac{2}{3}) = 0$.

    Strictly speaking we are confining our attention to continuous random variables. Using similar reasoning, we can readily find continuous examples. For example, suppose the probability density of X is given by $f(x) = \begin{cases} -\frac{1}{8}\sin\frac{x}{2}, & \text{for } -2\pi \le x \le 0 \\ \frac{1}{4}\sin x, & \text{for } 0 \le x \le \pi \end{cases}$.

11. Let X be a continuous random variable such that E$X = 0$. Let f be the probability density of X, and define g by $g(y) = f(y - \mu)$, where $\mu \in \mathbb{R}$ is a constant. Prove that g is a probability density and that if Y is a random variable with density g, then E$Y = \mu$. Both integrals may be computed by the substitution $x = y - \mu$.