

Sixth Problem Assignment

EECS 401

February 23, 2007

PROBLEM 1 (30 points) Joe Lucky plays the lottery on any given week with probability p , independently of whether he played on any other week. Each time he plays, he has a probability q of winning, again independently of everything else. During a fixed time period of n weeks, let X be the number of weeks that he played the lottery and Y be the number of weeks that he won.

- (a) What is the probability that he played the lottery on any particular week, given that he did not win on that week?

Solution Let L_i be the event that Joe played the lottery on week i , and let W_i be the event that he won on week i . The desired probability is

$$\Pr(L_i|W_i') = \frac{\Pr(W_i'|L_i) \Pr(L_i)}{\Pr(W_i'|L_i) \Pr(L_i) + \Pr(W_i'|L_i') \Pr(L_i')} = \frac{(1-q)p}{(1-q)p + 1 \cdot (1-p)} = \frac{p-pq}{1-pq}.$$

- (b) Find the conditional PMF $p_{Y|X}(y|x)$.

Solution Conditioned on X , the random variable Y is binomial

$$p_{Y|X}(y|x) = \begin{cases} \binom{x}{y} q^y (1-q)^{(x-y)}, & \text{if } 0 \leq y \leq x, \\ 0 & \text{otherwise} \end{cases}$$

- (c) Find the joint PMF $p_{X,Y}(x,y)$.

Solution Since X has a binomial PMF, we have

$$\begin{aligned} p_{X,Y}(x,y) &= p_{Y|X}(y|x) p_X(x) \\ &= \begin{cases} \binom{x}{y} q^y (1-q)^{(x-y)} \binom{n}{x} p^x (1-p)^{(n-x)}, & \text{if } 0 \leq y \leq x \leq n, \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

- (d) Find the marginal PMF $p_Y(y)$.

Hint: One possibility is to start with the answer to the previous part, but the algebra can be messy. However, if you think intuitively about the procedure that generates Y , you may be able to guess the answer.

Solution Using the result from part (c), we could compute the marginal p_Y using the formula

$$p_Y(y) = \sum_{x=y}^n p_{X,Y}(x,y),$$

but the algebra is messy. An easier method is as follows.

Let Y_i be a random variable that takes value 1 if Joe wins in week i and takes values 0 if Joe does not win in week i . Notice that Y_i is a Bernoulli random variable with

$$p_{Y_i}(y) = \begin{cases} pq & \text{if } y = 1 \\ 1 - pq & \text{if } y = 0 \end{cases}$$

Further, Y_i s are independent and $Y = Y_1 + \dots + Y_n$. Thus Y is a binomial random variable with

$$p_Y(y) = \begin{cases} \binom{n}{y} (pq)^y (1 - pq)^{(n-y)} & \text{if } 0 \leq y \leq n, \\ 0 & \text{otherwise} \end{cases}$$

- (e) Find the conditional PMF $p_{X|Y}(x|y)$. Do this algebraically using the preceding answers.

Solution We have

$$\begin{aligned} p_{X|Y}(x|y) &= \frac{p_{X,Y}(x,y)}{p_Y(y)} \\ &= \begin{cases} \frac{\binom{x}{y} q^y (1-q)^{(x-y)} \binom{n}{x} p^x (1-p)^{(n-x)}}{\binom{n}{y} (pq)^y (1-pq)^{(n-y)}} & 0 \leq y \leq x \leq n, \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

- (f) Rederive the answer to the preceding part by thinking as follows: for each one of the $n - Y$ weeks that he did not win, the answer to the first part of the problem should tell you something.

Solution From part (a), we know the probability $\Pr(L_i|W'_i)$ that Joe played the lottery on week i given that he did not win in that week. For each of the $n - y$ weeks when Joe did not win, there are $x - y$ weeks when he played. Thus, X conditioned on $Y = y$ is binomial with parameters $n - y$ and $\Pr(L_i|W'_i) = (p - pq)/(1 - pq)$:

$$p_{X|Y}(x|y) = \begin{cases} \binom{n-y}{x-y} \left(\frac{p-pq}{1-pq}\right)^{x-y} \left(1 - \frac{p-pq}{1-pq}\right)^{n-x} & 0 \leq y \leq x \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

After some algebraic manipulation, the answer to part (e) can be shown equal to the above.

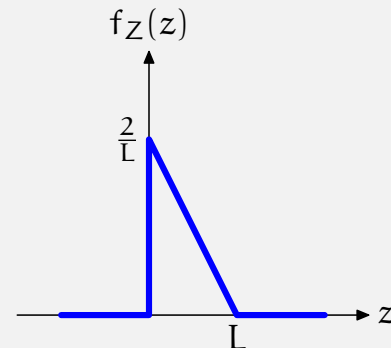
PROBLEM 2 (13 points) An ambulance travels back and forth, at a constant speed, along a road of length L . At a certain moment of time an accident occurs at a point uniformly distributed on the road. (That is, its distance from one of the fixed ends of the road is uniformly distributed over $(0, L)$.) Assuming that the ambulance's location at the moment of the accident is also uniformly distributed, compute, assuming independence, the distribution of its distance from the accident.

Solution Let X denote the position of the ambulance and Y denote the position of the accident. Both X and Y are uniformly distributed over $(0, L)$. Let $Z = |X - Y|$. Then we have the following:

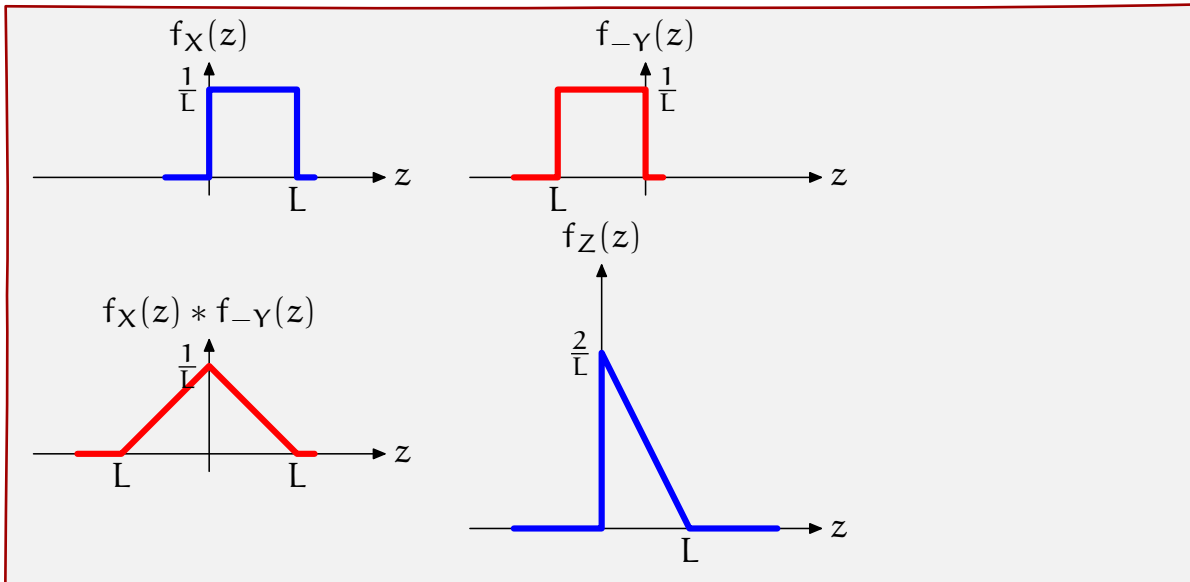
$$\begin{aligned} F_Z(z) &= \Pr(|X - Y| < z) = \Pr(-z < X - Y < z) = \int_0^L \Pr(y - z < X < y + z | Y = y) f_Y(y) dy \\ &= \int_0^L \Pr(y - z < X < y + z) f_Y(y) dy = \int_0^L [F_X(y + z) - F_X(y - z)] f_Y(y) dy \\ &= \int_0^z \frac{(y + z)}{L} \frac{1}{L} dy + \int_z^{L-z} \frac{2z}{L} \frac{1}{L} dy + \int_{L-z}^L \frac{(L + z - y)}{L} \frac{1}{L} dy \\ &= \frac{3z^2}{2L^2} + \frac{(L - 2z)(2z)}{L^2} + \frac{3z^2}{2L^2} = \boxed{\frac{2Lz - z^2}{L^2}} \end{aligned}$$

Thus to find the pdf $f_Z(z)$ of Z , we can differentiate $F_Z(z)$ to obtain:

$$f_Z(z) = \frac{dF_Z(z)}{dz} = \begin{cases} \frac{2(L-z)}{L^2}, & 0 < z < L \\ 0, & \text{otherwise} \end{cases}$$



This can also be *proved* graphically as follows.



PROBLEM 3 (10 points) Let X_1 , X_2 and X_3 be three independent, continuous random variables with the same distribution. Given that X_2 is smaller than X_3 , what is the conditional probability that X_1 is smaller than X_2 ?

Solution Consider the two random variables X_2 and X_3 . As they are independent and identically distributed, the events $\{X_2 > X_3\}$ and $\{X_3 > X_2\}$ are equally likely. Thus, we have: $\Pr(X_2 < X_3) = \frac{1}{2}$. Similarly,

$$\Pr(\{X_1 < X_2 < X_3\}) = \frac{1}{3!} = \frac{1}{6}.$$

Thus,

$$\Pr(X_1 < X_2 | X_2 < X_3) = \frac{\Pr(X_1 < X_2 < X_3)}{\Pr(X_2 < X_3)} = \frac{1/6}{1/2} = \boxed{\frac{1}{3}}$$

PROBLEM 4 (10 points) Random variables X and Y are described by the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 1, & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

and random variable Z is defined by $Z = XY$.

Determine the conditional second moment of Z , given that the equation $r^2 + xr + y = 0$ has real roots for r .

Solution Let R be the region where $r^2 + xr + y = 0$ has real roots. Then

$$\begin{aligned}
 R &= \{(x, y) : x^2 - 4y > 0\} \\
 E[z^2 \mid (x, y) \in R] &= \frac{\iint_R x^2 y^2 f_{X,Y}(x, y) dy dx}{\iint_R f_{X,Y}(x, y) dy dx} = \frac{\int_0^1 \int_0^{x^2/4} x^2 y^2 dy dx}{\int_0^1 \int_0^{x^2/4} dy dx} \\
 &= 12 \int_0^1 \frac{x^8}{192} dx = \boxed{\frac{1}{144}}
 \end{aligned}$$

PROBLEM 5 (27 points)

- (a) Variable X^* , the *standardized* random variable X , is given by $X^* = [X - E(X)]/\sigma_X$. Determine the expected value and variance of X^* .

Solution

$$\begin{aligned}
 E[X^*] &= E\left[\frac{X - E[X]}{\sigma_X}\right] = \frac{1}{\sigma_X} (E[X] - E[X]) = \boxed{0} \\
 \text{Var}(X^*) &= \frac{1}{\sigma_X^2} \text{Var}(X) = \boxed{1}
 \end{aligned}$$

- (b) The *correlation coefficient* ρ , or *normalized covariance*, for two random variables X and Y is defined to be

$$\rho_{XY} = E(X^*Y^*) = E\left[\left(\frac{X - E(X)}{\sigma_X}\right)\left(\frac{Y - E(Y)}{\sigma_Y}\right)\right]$$

Determine the numerical value of ρ_{XY} if:

- (i) $X = aY$.

Solution $X = aY \implies E[X] = aE[Y]$ and $\sigma_Y = |a|\sigma_X$. Thus, $\rho_{XY} = \text{sgn}(a)$, where sgn is the sign function, i.e. $\text{sgn}(a) = 1$ if $a > 0$, $\text{sgn}(a) = 0$ if $a = 0$, and $\text{sgn}(a) = -1$ if $a < 0$.

- (ii) $X = -aY$.

Solution $X = -aY \implies E[X] = -aE[Y]$ and $\sigma_Y = |a|\sigma_X$. Thus, $\rho_{XY} = -\text{sgn}(a)$

- (iii) X and Y are linearly independent.

Solution x and y are linearly independent. This means that $E[XY] = E[X]E[Y]$ and $E[X^*Y^*] = E[X^*]E[Y^*]$. Thus, $\rho_{XY} = 0$.

- (iv) X and Y are statistically independent.

Solution Statistical independence implies linear independence. Thus, $\rho_{XY} = 0$.

(v) $X = aY + b$.

Solution $X = aY + b \implies E[X] = aE[Y] + b$. Thus, $X - E[X] = a(Y - E[Y])$. and $\sigma_{(X-E[X])} = |a|\sigma_{(Y-E[Y])}$. Hence, $\rho_{XY} = \text{sgn}(a)$.

(c) For each performance of the experiment, the experimental value of random variable Y^* is to be approximated by cX^* . Prove that the value of constant c which minimizes the expected *mean square error*, $E[(Y^* - cX^*)^2]$, for this approximation is given by $c = \rho_{XY}$.

Solution

$$E[(Y^* - cX^*)^2] = E[Y^{*2}] + c^2E[X^{*2}] - 2cE[X^*Y^*] = 1 + c^2 - 2c\rho_{XY}$$

Taking derivative w.r.t c and setting it to zero we have

$$2c - 2\rho_{XY} = 0 \implies \boxed{c = \rho_{XY}}$$

PROBLEM 6 (10 points) Consider the following question and a purported solution. Either declare the solution to be correct, or explain the flaw.

Question: Let X and Y have the joint density

$$f_{X,Y}(x,y) = \begin{cases} 1, & x \in [0, 1] \text{ and } y \in [x, x+1] \\ 0, & \text{otherwise} \end{cases}$$

Find $f_X(x)$, $f_Y(y)$ and $f_{Y|X}(y|x)$. Are X and Y independent?

Solution:

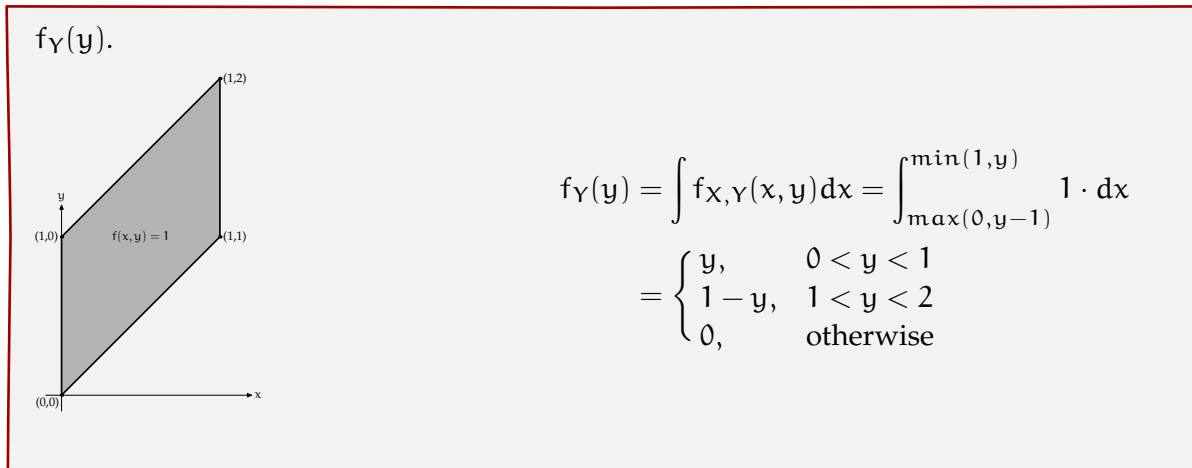
$$f_X(x) = \int f_{X,Y}(x,y) dy = \int_x^{x+1} 1 \cdot dy = 1$$

$$f_Y(y) = \int f_{X,Y}(x,y) dx = \int_0^1 1 \cdot dx = 1$$

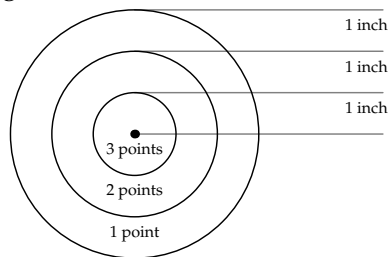
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{1} = 1.$$

Since $f_{Y|X}(y|x)$ does not depend on x , we have that X and Y are independent. Alternatively, X and Y are independent because $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$.

Solution The integral $f_Y(y)$ is evaluated incorrectly. A quick way to check this is that $y \in [0, 2]$ but $\int_0^2 f_Y(y) dy \neq 1$. This mistake is in the limits of integration while evaluating



Extra Credit (5 points) Each day Wyatt Uyrp shoots one “game” by firing at a target with the following dimensions and scores for each shot:



The score on any shot depends only on its distance from the center of the target

His pellet supply isn't too predictable, and the number of shots for any day's game is equally likely to be one, two, or three. Furthermore, Wyatt tires rapidly with each shot. Given that is the K th pellet in a particular game, the value of R (distance from target center to point of impact) for a pellet is a random variable with probability density function

$$f_{R|K}(r|k) = \begin{cases} \frac{1}{k}, & \text{if } 0 \leq r \leq k \\ 0, & \text{otherwise} \end{cases}$$

Given only that a particular pellet was used during a three-shot game, determine and sketch the probability density function for R , the distance from the target center to where it hit.

Solution By the law of total probability we have

$$f_R(r) = f_{R|K}(r|1)P_K(1) + f_{R|K}(r|2)P_K(2) + f_{R|K}(r|3)P_K(3)$$

We are given that

$$P_K(1) = P_K(2) = P_K(3) = \frac{1}{3}$$

Further,

$$f_{R|K}(r|1) = \begin{cases} 1, & 0 < r < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{R|K}(r|2) = \begin{cases} \frac{1}{2}, & 0 < r < 2 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{R|K}(r|3) = \begin{cases} \frac{1}{3}, & 0 < r < 3 \\ 0, & \text{otherwise} \end{cases}$$

Thus, we have

$$f_R(r) = \begin{cases} \frac{11}{18}, & 0 < r < 1 \\ \frac{5}{18}, & 1 < r < 2 \\ \frac{1}{9}, & 2 < r < 3 \end{cases}$$