14.30 Problem Set #3 Solutions
Due Tuesday, October 12, 2004
Guy Michaels

You must answer ALL questions, except the one marked OPTIONAL in order to fulfill the course requirement. For each student, the answers will be ordered by score and the top five answers will then comprise the final problem set grade.

Question 1:
Suppose that a random variable \( X_1 \) is distributed uniform \([0, 1]\), \( X_2 \) is distributed uniform \([0, 2]\) and \( X_3 \) is distributed uniform \([0, 3]\). Assume that they are all independent.

\( a. \) Calculate \( E [(X_1 - 2X_2 + X_3)] \)

**answer:** \( E [(X_1 - 2X_2 + X_3)] = E [X_1] - 2E [X_2] + E [X_3] = 2 (0.5 - 2 (1) + 1.5) = 0 \)

\( b. \) Calculate \( f_{X_1,X_2,X_3} (x_1, x_2, x_3) \)

**answer:** \( f_{X_1,X_2,X_3} (x_1, x_2, x_3) = \begin{cases} \frac{1}{6} & \text{if } x_1 \in [0, 1], x_2 \in [0, 2], x_3 \in [0, 3] \\ 0 & \text{otherwise} \end{cases} \)

\( c. \) Use your result from part \( b. \) to calculate \( E [(X_1 - 2X_2 + X_3)^2] \)

**answer:** \( Define \( g (X_1, X_2, X_3) = (X_1 - 2X_2 + X_3)^2 \). \( E [(X_1 - 2X_2 + X_3)^2] = E [g (X_1, X_2, X_3)] = \int_0^1 \int_0^2 \int_0^3 \frac{1}{6} (x_1 - 2x_2 + x_3)^2 \, dx_1 \, dx_2 \, dx_3 = \frac{1}{6} \int_0^2 \int_0^3 \left[ \frac{1}{3} + 4x_2^2 + x_3^2 - 2x_2 - 4x_2x_3 + x_3 \right] \, dx_2 \, dx_3 = \frac{1}{6} \int_0^3 (\frac{22}{3} + 2x_3^2 - 6x_3) \, dx_3 = \frac{1}{6} (22 + 18 - 27) = \frac{13}{6} \)

\( d. \) Use your result from parts \( a. \) and \( c. \) to calculate \( Var [(X_1 - 2X_2 + X_3)] \)

**answer:** \( Var [Z] = E [Z^2] - (E [Z])^2 \), so: \( Var [(X_1 - 2X_2 + X_3)] = E [(X_1 - 2X_2 + X_3)^2] - (E [(X_1 - 2X_2 + X_3)])^2 = \frac{13}{6} - 0^2 = \frac{13}{6} \)

\( e. \) Calculate \( E [(X_1 + X_2) | X_3] \)

**answer:** since \( X_1 \) and \( X_2 \) are independent of \( X_3 \) so is their sum. Therefore: \( E [(X_1 + X_2) | X_3] = E [X_1 + X_2] = E [X_1] + E [X_2] = \frac{3}{2} \)

Question 2:
Suppose that \( X, Y \) and \( e \) are random variables with \( Y = \alpha + \beta X + e \). Assume \( E (X) = 0, Var (X) = \sigma^2; e \) has a uniform distribution over \([ -\frac{1}{2}, \frac{1}{2} ] \) and \( Cov(X, e) = 0 \).

\( a. \) What is the distribution of \( Y | X? \)

**answer:** The distribution of \( Y | X \) is the distribution of \( Y \), where \( e \) is the only random variable and \( \alpha + \beta X \) is fixed, say \( C \). The distribution of \( Y = C + e \) is thus uniform \([ C - \frac{1}{2}, C + \frac{1}{2} ] \)
b. What is \( E[Y|X] \)? What is \( Var[Y|X] \)?

**Answer:** If \( U \) is distributed uniform on \([a,b] \) then: \( f_u(u) = \frac{1}{b-a}, \) \( E[U] = \frac{a+b}{2}, Var[U] = \frac{(b-a)^2}{12}. \) Therefore: \( E[Y|X] = \frac{(C-\beta)}{2} = C = \alpha + \beta X, Var[Y|X] = \frac{(C+\beta)-(C-\beta)^2}{12} = \frac{1}{12} \)

c. What is \( E[Y] \)? What is \( Var[Y] \)?

**Answer:** \( E[Y] = E[\alpha + \beta X + e] = \alpha + \beta E[X] + E[e] = \alpha. \) \( Var[Y] = Var[\beta X + e] = \beta^2 Var[X] + Var[e] + 2\beta Cov(X,e) = \beta^2 \sigma^2 + \frac{1}{12} \)

**Question 3:**
Suppose that you are considering undertaking a project where the revenue, \( X_1 \), has a distribution: \( f_{X_1}(x_1) = \begin{cases} \frac{3}{4} & \text{if } x_1 = 0 \\ \frac{1}{4} & \text{if } x_1 = 8 \end{cases} \). The project cost, \( Y_1 \), is distributed: \( f_{Y_1}(y_1) = \begin{cases} \frac{1}{2} & \text{if } y_1 \in \{1, \frac{3}{4} \} \\ 0 & \text{otherwise} \end{cases} \). Assume that the revenue and the cost are independent.

a. What is the expected profit (revenue minus cost) of the project? Would your answer be different if \( X_1 \) and \( Y_1 \) were not independent?

**Answer:** \( E[X_1 - Y_1] = E[X_1] - E[Y_1] = \left( \frac{3}{4} \right) 8 - \frac{1}{2} \left( 1 + \frac{3}{4} \right) = \frac{3}{4} \). This is true regardless of independence.

b. What is the probability that the project makes a positive profit? Would your answer be different if \( X_1 \) and \( Y_1 \) were not independent?

**Answer:** \( \Pr[X_1 - Y_1 > 0] = \Pr[X_1 = 8] = \frac{1}{4} \). This is also true regardless of independence.

c. Suppose now that a friend can undertake a different project where the revenue \( (X_2) \) is distributed identically to \( X_1 \) and the the cost \( (Y_2) \) is distributed identically to \( Y_1 \). All the variables \( X_1, Y_1, X_2, Y_2 \) are independent. Before you undertake either project your friend suggests that you equally share the revenue and costs of both projects. What is the expected profit to accepting the suggestion relative to not taking up any project? What is the probability that you make a positive profit if you accept your friend’s suggestion?

**Answer:** \( E\left[ \frac{1}{4}(X_1 - Y_1 + X_2 - Y_2) \right] = \frac{3}{4} \). The probability of making a positive profit are now: \( \Pr\left[\frac{1}{4}(X_1 - Y_1 + X_2 - Y_2) > 0\right] = \Pr[X_1 = 8 \cup X_2 = 8] = \Pr[X_1 = 8] + \Pr[X_2 = 8] - \Pr[X_1 = X_2 = 8] = \frac{1}{4} + \frac{1}{4} - \left( \frac{1}{4} \right) \left( \frac{1}{4} \right) = \frac{7}{16} \)

**Question 4:**
Let \( X \) and \( Y \) be two random variables, where \( f_{X,Y}(-1,1) = \frac{1}{2}, f_{X,Y}(0,2) = \frac{1}{4}, f_{X,Y}(2,1) = \frac{1}{4}. \)

a. What is the correlation between \( X \) and \( Y \)?
answer: \( E[X] = 0, E[Y] = \frac{5}{4} \Rightarrow E[X] E[Y] = 0 \). Now: \( E[XY] = \frac{1}{2} (-1) + \frac{1}{4} (2) = 0 \Rightarrow Cov(X,Y) = E[XY] - E[X] E[Y] = 0 - 0 = 0 \Rightarrow Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} = 0 \) so \( X \) and \( Y \) are uncorrelated.

b. Are \( X \) and \( Y \) independent?

answer: \( X \) and \( Y \) are not independent, for example: \( f_{X,Y}(0,1) = 0 \neq \left( \frac{1}{4} \right) \left( \frac{3}{4} \right) = f_X(0) f_Y(1) \).

c. What do you conclude about the relationship between correlation and independence?

answer: You saw in class that independence implies uncorrelatedness. This exercise proves that two uncorrelated random variables are not necessarily independent.

Question 5:

Suppose that \( X \) is a discrete random variable \( f_X(x) = \begin{cases} \frac{3}{10} & \text{if } x = 10 \\ \frac{1}{10} & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases} \) and \( Y \) is a continuous random variable distributed uniformly over \([-2,8]\).

a. Which variable has a higher mean?

answer: \( E[X] = \frac{3}{10} \cdot 10 = 3 \). \( E[Y] = \frac{-2+8}{2} = 3 \). So both variables have the same mean.

b. Which variable has a higher median?

answer: The median of \( X \) is 0 and the median of \( Y \) is 3. Therefore \( Y \) has a higher median.

Question 6:

a. Assume that \( X \) has mean 2 and variance \( \frac{4}{3} \). Use the Chebyshev inequality to calculate an upper bound on the probability is outside the interval \((0.5, 3.5)\).

answer: The Chebyshev inequality states that \( \Pr \left( |X - E[X]| \geq t \right) \leq \frac{\text{Var}[X]}{t^2} \) for \( t > 0 \). Here: \( E[X] = 2, \text{Var}[X] = \frac{4}{3}, t = \frac{3}{2} \) so: \( \Pr (X \notin (0.5, 3.5)) = \Pr (|X - 2| \geq \frac{3}{2}) \leq \frac{(\frac{4}{3})}{(\frac{3}{2})^2} = \frac{16}{27} \).

b. Now assume that \( X \sim U[0, 4] \). Calculate the probability that \( X \) is outside the interval \((0.5, 3.5)\). Is it higher or lower than the answer you got in part a.?

answer: \( \Pr (X \notin (0.5, 3.5)) = \frac{1}{4} < \frac{16}{27} \) so we get a lower result than in part a.

c. Based on your answers to parts a. and b., comment on the usefulness of the Chebyshev inequality.
The Chebyshev inequality is very useful for evaluating distributions for which you only know the mean and the median, but not the actual distribution. If you know the actual distribution you can get a more precise answer. But this is only because you are using additional information.

Question 7:

a. Let $X$ be a continuous random variable with a pdf $f_X(x)$ and let $a,b$ and $c$ be constants. Prove that for any functions $g_1(x)$ and $g_2(x)$ whose expectations exist, $E(\sum_{k=0}^{\infty} X_g(x)+g_2(x)+c) = aE [g_1 (X)] + bE [g_2 (X)] + c$.

b. According to the model, what is the expected time until the price in the market reaches its equilibrium level (in months) is described using the pmf: An economist has constructed a model where the time it takes for the price in a given market to reach its equilibrium level (in months) is described using the pmf:

$Pr \left[ X=k \right] = p^k (1-p), \quad k \in \mathbb{N}$

where

$X \sim U [0,1]$. Here: $E(g_1(X)) = \int_0^1 x^2dx = \frac{1}{3}$; but: $g_1(E[X]) = \left( \int_0^1 xdx \right)^2 = \frac{1}{4}$. Thus:

$E(g_1(X)) > g_1(E[X])$. More generally, Jensen’s Inequality states that for a strictly convex function $g_1(E[X])$ (such as $g_1(x) = x^2$): $E(g_1(X)) > g_1(E[X])$. For a strictly concave function $g_1(E[X]) : E(g_1(X)) < g_1(E[X])$.

c. prove that $E[X] = E[E[X|Y]]$ for any two continuous random variables $X$ and $Y$.

Question 8 (optional):

An economist has constructed a model where the time it takes for the price in a given market to reach its equilibrium level (in months) is described using the pmf: $Pr \left[ X=k \right] = p^k (1-p)$, where $p \in (0,1)$ and $k$ is a non-negative integer.

a. Prove that $\sum_{k=0}^{\infty} Pr \left[ X=k \right] = 1$.

b. According to the model, what is the expected time until the price in the market reaches its equilibrium level?

$E[X] = \lim_{n \to \infty} \sum_{k=0}^{n} p^k (1-p) = p (1-p) \lim_{n \to \infty} \sum_{k=0}^{n} p^{k-1} = p (1-p) \lim_{n \to \infty} \sum_{k=0}^{n} \frac{\partial(p^k)}{\partial k} = p (1-p) \frac{\partial}{\partial k} \left[ \lim_{n \to \infty} \sum_{k=0}^{n} p^k \right] = p (1-p) \frac{\partial}{\partial k} \left[ \frac{1}{1-p} \right] = p (1-p) \frac{1}{(1-p)^2} = \frac{p}{(1-p)}$. 
