Stat 310 Homework 5 Key

Chapter 4, problems 51, 58, 61, 67, 69, 70, 71, 76, 85, 87. Due 10/7/99.

4.50 BONUS WORKED PROBLEM! I LIKED THIS ONE, SO I WENT AHEAD AND DID IT. THIS ONE WAS NOT ASSIGNED. Let $S = \sum_{k=1}^{n} X_k$, where the X_k are as in Problem 49. Find the covariance and correlation of S and T.

Ok, from problem 49, we have that $T = \sum_{k=1}^{n} kX_k$, where the X_k are independent random variables with means μ and variances σ^2. To find the covariance of S and T, we need to find $E(S)$, $E(T)$, and $E(ST)$. The first of these follows straight from linearity:

$$\begin{aligned} E(S) &= E(\sum_{k=1}^{n} X_k) \\ &= \sum_{k=1}^{n} E(X_k) \\ &= n\mu. \end{aligned}$$

The next is just about as straightforward:

$$\begin{aligned} E(T) &= E(\sum_{k=1}^{n} kX_k) \\ &= \sum_{k=1}^{n} kE(X_k) \\ &= \frac{n(n+1)}{2}\mu. \end{aligned}$$

Now for the big one, the expectation of the product.

$$\begin{aligned} E(ST) &= E((\sum_{k=1}^{n} X_k)(\sum_{j=1}^{n} jX_j)) \\ &= \sum_{k=1}^{n}\sum_{j=1}^{n} jE(X_kX_j) \end{aligned}$$

Now, if the j and k are different, $E(X_kX_j) = E(X_k)E(X_j) = \mu^2$ by independence. If the indices are the same, $E(X_kX_k) = \sigma^2 + \mu^2$. So, every term contributes a μ^2, but only those on the main diagonal contribute σ^2 terms. Thus,

$$\begin{aligned} E(ST) &= \sum_{k=1}^{n}\sum_{j=1}^{n} j\mu^2 + \sum_{k=1}^{n} k\sigma^2 \\ &= n\frac{n(n+1)}{2}\mu^2 + \frac{n(n+1)}{2}\sigma^2 \end{aligned}$$

and the covariance is

$$\begin{aligned} Cov(S,T) &= E(ST) - E(S)E(T) \\ &= \frac{n(n+1)}{2}\sigma^2. \end{aligned}$$

This can actually be found by a heuristic "matching" argument which uses the linearity of the expectation to show that we have

$$
\begin{aligned}
Cov(S,T) &= E(ST) - E(S)E(T) \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n} j\,[E(X_k X_j) - E(X_k)E(X_j)]
\end{aligned}
$$

and noting that all of the terms where the indices do not match contribute zero, as the X's are independent, and the cases where the indices do match contribute $j\sigma^2$, so

$$
\begin{aligned}
Cov(S,T) &= \sum_{j=1}^{n} j\sigma^2 \\
&= \frac{n(n+1)}{2}\sigma^2
\end{aligned}
$$

just as before. Now, for the correlation of S and T, we need the variances of S and T. Here,

$$
\begin{aligned}
V(S) &= V(\sum_{k=1}^{n} X_k) \\
&= \sum_{k=1}^{n} V(X_k) \\
&= n\sigma^2, \\
V(T) &= V(\sum_{k=1}^{n} kX_k) \\
&= \sum_{k=1}^{n} k^2 V(X_k) \\
&= \frac{n(n+1)(2n+1)}{6}\sigma^2.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
Corr(S,T) &= \frac{Cov(S,T)}{\sqrt{Var(S)}\sqrt{Var(T)}} \\
&= \frac{\frac{n(n+1)}{2}\sigma^2}{\sqrt{n\sigma^2}\sqrt{\frac{n(n+1)(2n+1)}{6}\sigma^2}} \\
&= \frac{\frac{n+1}{2}}{\sqrt{\frac{(n+1)(2n+1)}{6}}} \\
&= \sqrt{\frac{3(n+1)}{2(2n+1)}}.
\end{aligned}
$$

4.51 If X and Y are independent random variables, find $E(XY)$ in terms of the means and variances of X and Y.

$$
\begin{aligned}
V(XY) &= E((XY)^2) - E(XY)^2 \\
&= E(X^2)E(Y^2) - E(X)^2E(Y)^2 \\
&= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 \\
&= \sigma_X^2\sigma_Y^2 + \mu_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2.
\end{aligned}
$$

Up above, we rearranged the standard formula for the variance to isolate $E(X^2)$ as follows:

$$
\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
E(X^2) &= V(X) + E(X)^2.
\end{aligned}
$$

4.58. Let X and Y be jointly distributed random variables with correlation ρ_{XY}; define the *standardized* random variables $\tilde{X}$ and $\tilde{Y}$ as $\tilde{X} = (X - E(X))/\sqrt{Var(X)}$ and $\tilde{Y} = (Y - E(Y))/\sqrt{Var(Y)}$. Show that $Cov(\tilde{X}, \tilde{Y}) = \rho_{XY}$.

$$
\begin{aligned}
Cov(\tilde{X}, \tilde{Y}) &= E(\tilde{X}\tilde{Y}) - E(\tilde{X})E(\tilde{Y}) \\
&= E\left(\frac{X - E(X)}{\sqrt{Var(X)}}\frac{Y - E(Y)}{\sqrt{Var(Y)}}\right) - E\left(\frac{X - E(X)}{\sqrt{Var(X)}}\right) E\left(\frac{Y - E(Y)}{\sqrt{Var(Y)}}\right) \\
&= \frac{1}{\sqrt{Var(X)}\sqrt{Var(Y)}} * \\
&\quad [E((X - E(X))(Y - E(Y))) - E(X - E(X))E(Y - E(Y))] \\
&= \frac{1}{\sqrt{Var(X)}\sqrt{Var(Y)}} [E(XY) - E(X)E(Y) - 0] \\
&= \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \rho_{XY}.
\end{aligned}
$$

4.61. A random rectangle is formed in the following way: The base, X, is chosen to be a uniform $[0, 1]$ random variable and after having generated the base, the height is chosen to be uniform on $[0, X]$. Use the law of total expectation, Theorem A of Section 4.4.1, to find the expected circumference and area of the rectangle.

Let H be the height of the rectangle. The circumference of the rectangle is $2(X + H)$, so we want the expected value of this quantity. The expected value of H is a lot easier to find if we condition on the value of X. Before we get started, a side note on the expectation of a uniform $[a, b]$ random variable (call it Z):

$$
E(Z) = \int_a^b z f_Z dz = \int_a^b \frac{z}{b - a} = \frac{a + b}{2}.
$$

We'll be using the expected value of a uniform distribution quite a bit below.

$$
\begin{aligned}
E(2(X + H)) &= E[E(2(X + H)|X)] && \text{law of total E} \\
&= E[2X + 2E(H|X)] && \text{note, } E(X|X) = X \\
&= E\left[2X + 2\frac{X}{2}\right] = \frac{3}{2} && \text{using expected values of uniforms.}
\end{aligned}
$$

Now for the area. The area is simply XH, so

$$\begin{aligned}
E(XH) &= E[E(XH|X)] \\
&= E[XE(H|X)] \\
&= \frac{1}{2}E(X^2) \\
&= \frac{1}{2}\int_0^1 x^2 dx = \frac{1}{6}.
\end{aligned}$$

4.67. A fair coin is tossed n times, and the number of heads, N, is counted. The coin is then tossed N more times. Find the expected total number of heads generated by this process.

Let X denote the number of heads in the second stage of the process, so that we want $E(N+X)$. To do this, we need to identify the distributions involved. In the first stage, the number of heads is binomially distributed, with parameters n and $1/2$ (we're told that the coin is fair). The expected value of a binomial random variable we found in class as n times the expected value of a Bernoulli random variable, p, or np. In the second stage, the number of heads is again (conditionally) binomial with parameters N and $1/2$.

$$\begin{aligned}
E(N+X) &= E[E(N+X|N)] \\
&= E[N+E(X|N)] \\
&= \frac{3}{2}E(N) = \frac{3}{4}n.
\end{aligned}$$

In complete generality, if the coin is not necessarily fair, then we work with p, getting

$$\begin{aligned}
E(N+X) &= E[E(N+X|N)] \\
&= E[N+E(X|N)] \\
&= E(N+Np) \\
&= (1+p)E(N) = np(1+p).
\end{aligned}$$

4.69. Let T be an exponential random variable, and conditional on T, let U be uniform on $[0,T]$. Find the unconditional mean and variance of U.

Ok, we're going to tackle this using the law of total expectation again, finding both the first and second moments of U.

$$\begin{aligned}
E(U) &= E[E(U|T)] \\
&= \frac{1}{2}E(T) \\
&= \frac{1}{2}\int_0^\infty \lambda t e^{-\lambda t} \\
&= \frac{1}{2\lambda}.
\end{aligned}$$

$$
\begin{aligned}
E(U^2) &= E[E(U^2|T)] \\
&= E\left[\int_0^T u^2 \frac{1}{T} du\right] \\
&= E\left[\frac{1}{T}\int_0^T u^2 du\right] \\
&= \frac{1}{3}E(T^2) \\
&= \frac{1}{3}\int_0^\infty \lambda t^2 e^{-\lambda t} \\
&= \frac{1}{3}\frac{2}{\lambda^2} = \frac{2}{3\lambda^2} \\
V(U) &= E(U^2) - E(U)^2 \\
&= \frac{2}{3\lambda^2} - 14\lambda^2 = \frac{5}{12\lambda^2}.
\end{aligned}
$$

4.70 Let the point (X, Y) be uniformly distributed over the half disk $x^2 + y^2 \leq 1$, where $y \geq 0$. If you observe X, what is the best prediction for Y? If you observe Y, what is the best prediction for X? For both questions, "best" means having the minimum mean squared error.

Ok, we showed in class that if we are given X, then our best guess, c, as to the value of Y is given by the minimizer of

$$
\begin{aligned}
E((Y-c)^2|X) &= V((Y-c)|X) + E((Y-c)|X)^2 \\
&= V(Y) + [E(Y|X) - c]^2.
\end{aligned}
$$

As the first term on the right does not depend on c, we choose it to minimize just the second term and get that $c = E(Y|X)$. So, this is what the problem is asking for. For these, it would be useful to have the marginal distributions of X and Y; we note that $f_{XY} = 2/\pi$ over the range where it is nonzero.

$$
\begin{aligned}
f_X(x) &= \int_0^{\sqrt{1-x^2}} \frac{2}{\pi} dy \\
&= \frac{2\sqrt{1-x^2}}{\pi}, \quad -1 \leq x \leq 1 \\
f_Y(y) &= \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{2}{\pi} dx \\
&= \frac{4\sqrt{1-y^2}}{\pi}, \quad 0 \leq y \leq 1.
\end{aligned}
$$

Now, for the conditional expectations:

$$
\begin{aligned}
E(Y|X) &= \int y f_{Y|X}(y) dy \\
&= \int y \frac{f_{XY}}{f_X} dy
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^{\sqrt{1-x^2}} y\frac{2}{\pi}\frac{\pi}{2\sqrt{1-x^2}}dy \\
&= \frac{1}{\sqrt{1-x^2}} * \frac{y^2}{2}\bigg|_0^{\sqrt{1-x^2}} = \frac{\sqrt{1-x^2}}{2}. \\
E(X|Y) &= \int x f_{X|Y}(x)dx \\
&= \int x\frac{f_{XY}}{f_Y}dx \\
&= \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x\frac{2}{\pi}\frac{\pi}{4\sqrt{1-y^2}}dx \\
&= \frac{1}{2\sqrt{1-y^2}} * \frac{x^2}{2}\bigg|_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} = 0.
\end{aligned}
$$

If you take a look at a picture, these correspond to the centers of the appropriate vertical and horizontal slice of the half-disk, so that makes sense.

4.71. Let X and Y have the joint density

$$f(x,y) = e^{-y}, \quad 0 \le x \le y.$$

a) Find $Cov(X,Y)$ and the correlation of X and Y. It is a bit easier to do this by first finding the marginal densities of X and Y.

$$
\begin{aligned}
f_X(x) &= \int_x^\infty e^{-y}dy \\
&= e^{-x}, \quad 0 \le x \le \infty. \quad \text{exponential, } \lambda = 1 \\
f_Y(y) &= \int_0^y e^{-y}dx \\
&= e^{-y}\int_0^y dx \\
&= ye^{-y}, \quad 0 \le y \le \infty. \quad \text{gamma, } \lambda = 1, \alpha = 2.
\end{aligned}
$$

Now, for the covariance, we need $E(XY)$, $E(X)$, and $E(Y)$.

$$
\begin{aligned}
E(XY) &= \int_0^\infty \int_0^y xye^{-y}dxdy \\
&= \int_0^\infty ye^{-y}\left[\int_0^y xdx\right]dy \\
&= \int_0^\infty \frac{1}{2}y^3e^{-y}dy \\
&= \frac{1}{2}\int_0^\infty y^3e^{-y}dy = \frac{1}{2}\Gamma(4) = 3. \\
E(X) &= \int_0^\infty xe^{-x}dx = \Gamma(2) = 1. \\
E(Y) &= \int_0^\infty y^2e^{-y}dy = \Gamma(3) = 2. \\
Cov(X,Y) &= E(XY) - E(X)E(Y) = 1.
\end{aligned}
$$

For the correlation, we need the variances of X and Y, or more simply just the second moments.

$$\begin{aligned}
E(X^2) &= \int_0^\infty x^2 e^{-x} dx = \Gamma(3) = 2. \\
E(Y^2) &= \int_0^\infty y^3 e^{-y} dy = \Gamma(4) = 6. \\
V(X) &= E(X^2) - E(X)^2 = 1 \\
V(Y) &= E(Y^2) - E(Y)^2 = 2 \\
Corr(X, Y) &= \frac{1}{\sqrt{2}}.
\end{aligned}$$

b) Find $E(X|Y = y)$ and $E(Y|X = x)$.

$$\begin{aligned}
E(X|Y = y) &= \int x f_{X|Y=y}(x) dx \\
&= \int x \frac{f_{XY}(x, y)}{f_Y(y)} dx \\
&= \int_0^y x \frac{e^{-y}}{ye^{-y}} dx \\
&= \frac{1}{y} \int_0^y x dx = \frac{1}{y} \frac{x^2}{2}\bigg|_0^y = \frac{y}{2}. \\
E(Y|X = x) &= \int y f_{Y|X=x}(y) dy \\
&= \int y \frac{f_{XY}(x, y)}{f_X(x)} dy \\
&= \int_x^\infty y \frac{e^{-y}}{e^{-x}} dy \\
&= e^x \int_x^\infty ye^{-y} dy \\
&= e^x \left[-ye^{-y}|_x^\infty + \int_x^\infty e^{-y} dy \right] \\
&= e^x \left[xe^{-x} + e^{-x} \right] = x + 1.
\end{aligned}$$

c) Find the density functions of the random variables $E(X|Y)$ and $E(Y|X)$.

$$\begin{aligned}
F_{E(X|Y)}(c) &= P\left(E(X|Y) < c\right) \\
&= P\left(\frac{Y}{2} < c\right) \\
&= P(Y < 2c) = F_Y(2c) \\
f_{E(X|Y)}(c) &= \frac{\partial}{\partial c} F_Y(2c) = 2f_Y(2c) \\
&= 4ce^{-2c}, \quad 0 < c < \infty \quad \text{gamma}, \lambda = 2, \alpha = 1. \\
F_{E(Y|X)}(c) &= P\left(E(Y|X) < c\right) \\
&= P\left(X + 1 < c\right) \\
&= P(X < c - 1) = F_X(c - 1)
\end{aligned}$$

$$
\begin{aligned}
f_{E(X|Y)}(c) &= \frac{\partial}{\partial c} F_X(c-1) = f_X(c-1) \\
&= e^{-(c-1)}, \quad 1 < c < \infty \quad \text{shifted exponential, } \lambda = 1.
\end{aligned}
$$

Note that this last distribution is new - a continuous distribution offset by a constant.

4.76. Use the result of problem 75 to find the mgf of a binomial random variable and its mean and variance.

Ok, problem 75 asks us to find the mgf of a Bernoulli random variable; given this we can find the mgf of a binomial as a binomial can be expressed as a sum of independent Bernoulli trials. For the Bernoulli,

X	0	1
p_X	q	p

where $q = 1 - p$. The moment generating function is simply

$$M_X(t) = qe^0 + pe^t = q + pe^t.$$

Let Y be a binomial random variable with paramters n and p. Then we can write $Y = X_1 + \ldots + X_n$, and

$$M_Y(t) = M_X(t)^n = (q + pe^t)^n.$$

Now, for the mean and variance,

$$
\begin{aligned}
M_Y'(t) &= npe^t(q + pe^t)^{n-1}, \\
M_Y'(0) &= np \\
M_Y''(t) &= n(n-1)p^2e^{2t}(q + pe^t)^{n-2} + npe^t(q + pe^t)^{n-1} \\
M_Y''(0) &= n(n-1)p^2 + np \\
\mu_Y &= M_Y'(0) = np, \\
\sigma_Y^2 &= M_Y''(0) - (M_Y'(0))^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)
\end{aligned}
$$

4.85. Use the mgf to show that if X follows an exponential distribution, cX $(c > 0)$ does also.

The moment generating function for an exponential random variable is

$$E(e^{tX}) = \int_0^\infty \lambda e^{-(\lambda - t)x} dx = \frac{\lambda}{\lambda - t}.$$

We showed this in class on Thursday. In general, the moment generating function of $Y = cX$ is

$$E(e^{tY}) = E(e^{tcX}) = M_X(ct);$$

this was also shown in class. Combining the two, the moment generating function of $Y = cX$ when X is exponential is

$$M_Y(t) = \frac{\lambda}{\lambda - ct} = \frac{\frac{\lambda}{c}}{\frac{\lambda}{c} - t}$$

which is the moment generating function of an exponential random variable with parameter λ/c. Hence, Y is also an exponential random variable (albeit with a different parameter value than X).

4.87. Find the distribution of a geometric sum of exponential random variables by using moment generating functions.

Ok, we want to find the distribution of

$$Y = \sum_{i=1}^{N} X_i,$$

where N is a geometric random variable and each X_i is exponential. We will assume that the X_i's are all iid - independent and identically distributed. To find the distribution of the sum, we'll try to find the moment generating function of Y. Now, the unconditional distribution of Y is not all that straightforward, but if we condition on the value of N things get a good deal simpler. Indeed, the distribution of Y given $N = n$ is gamma, with parameters n and λ. We'll try to incorporate this by using the law of total expectation.

$$\begin{aligned}
E(e^{tY}) &= E[E(e^{tY}|N)] \\
&= E\left[\left(\frac{\lambda}{\lambda - t}\right)^N\right] \\
&= \sum_{k=1}^{\infty} \left(\frac{\lambda}{\lambda - t}\right)^k pq^{k-1} \\
&= \frac{p\lambda}{\lambda - t} \sum_{k=1}^{\infty} \left(\frac{q\lambda}{\lambda - t}\right)^{k-1} \\
&= \frac{p\lambda}{\lambda - t} \sum_{j=0}^{\infty} \left(\frac{q\lambda}{\lambda - t}\right)^{j} \\
&= \frac{\frac{p\lambda}{\lambda - t}}{1 - \frac{q\lambda}{\lambda - t}} \\
&= \frac{p\lambda}{\lambda - t - q\lambda} \\
&= \frac{p\lambda}{p\lambda - t}.
\end{aligned}$$

This is the moment generating function of an exponential random variable, with parameter $p\lambda$; this is the distribution of the sum.