NANYANG TECHNOLOGICAL UNIVERSITY

SPMS/DIVISION OF MATHEMATICAL SCIENCES

2016/17 Semester 1 MH2500 Probability and Introduction to Statistics Tutorial 11

For the tutorial on 3 November, let us discuss

- Ex. 4.7. 70, 73, 77, 80, 85, 96.

Ex. 4.7.70 If X and Y are independent, show that $E(X|Y=y)=E(X)$.

[Solution:] For discrete case,

$$\begin{aligned}
E(X|Y=y) &= \sum_x x p_{X|Y}(x|y) \\
&= \sum_x \frac{xP(X=x \text{ and } Y=y)}{P(Y=y)} \\
&= \sum_x \frac{xP(X=x)P(Y=y)}{P(Y=y)} \qquad \text{(since } X \text{ and } Y \text{ are independent)} \\
&= \sum_x xP(X=x) = E(X).
\end{aligned}$$

For continuous case,

$$\begin{aligned}
E(X|Y=y) &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx \\
&= \int_{-\infty}^{\infty} \frac{x f_{X,Y}(x,y)}{f_Y(y)}dx \\
&= \int_{-\infty}^{\infty} \frac{x f_X(x) f_Y(y)}{f_Y(y)}dx \qquad \text{(since } X \text{ and } Y \text{ are independent)} \\
&= \int_{-\infty}^{\infty} x f_X(x)dx = E(X).
\end{aligned}$$

Ex. 4.7.73. 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.

[Solution:] Method 1:

Let N denote the number of heads in the first n tosses and let H denote the number of heads in the N tosses. Then N is binomial with parameters n and p and $H|N$ is also binomial with parameters N and p. Therefore, $E(H|N)=Np$ and the expected number is

$$\begin{aligned}
E(N+H) &= E(N)+E(H) \\
&= np + E(E(H|N) = np + E(Np) = np + pE(N) = np + np^2 = np(1+p).
\end{aligned}$$

Since a fair coin is used, $p=1/2$ and so $E(X)=n\frac{1}{2}\frac{3}{2}=\frac{3n}{4}$.

Method 2: We compute directly. First note that

$$\sum_{l=0}^{k} l\binom{k}{l}p^l(1-p)^{k-l} = \sum_{l=1}^{k}\binom{k}{l}p^l(1-p)^{k-l} = kp\sum_{l=0}^{k-1}\binom{k-1}{l-1}p^{l-1}(1-p)^{k-l} = kp.$$

Let k represents the number of heads in the first n tosses and let l represent the number of heads in the next k tosses. Then computing directly,

$$\begin{aligned}
E(X) &= \sum_{k=0}^{n}\sum_{l=0}^{k}(k+l)\binom{n}{k}\binom{k}{l}p^k(1-p)^{n-k}p^l(1-p)^{k-l} \\
&= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}\sum_{l=0}^{k}\binom{k}{l}p^l(1-p)^{k-l} + \sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}\sum_{l=0}^{k} l\binom{k}{l}p^l(1-p)^{k-l} \\
&= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} + p\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}k \\
&= np + np^2 \\
&= np(1+p).
\end{aligned}$$

Ex. 4.7.77. Let X and Y have the joint density

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

a. Find $\mathrm{Cov}(X,Y)$ and the correlation of X and Y.
b. Find $E(X|Y=y)$ and $E(Y|X=x)$.
c. Find the density functions of the random variables $E(X|Y)$ and $E(Y|X)$.

[Solution:]

a. First, the marginal densities are

$$\begin{aligned}
f_X(x) &= \int_x^\infty e^{-y}dy = [-e^{-y}]_x^\infty = e^{-x}, &&(x \ge 0) \\
f_Y(y) &= \int_0^y e^{-y}dx = [xe^{-y}]_0^y = ye^{-y} &&(y \ge 0).
\end{aligned}$$

Next, we compute the expected values.

$$\begin{aligned}
E(X) &= \int_0^\infty xe^{-x}dx &= [-xe^{-x}]_0^\infty + \int_0^\infty e^{-x}dx &= 0 + [-e^{-x}]_0^\infty = 1 \\
E(Y) &= \int_0^\infty y^2e^{-y}dy &= [-y^2e^{-y}]_0^\infty + \int_0^\infty 2ye^{-y}dy &= 2.
\end{aligned}$$

Similarly, we can show that $E(X^2) = 2$ and $E(Y^2) = 6$ and so $\mathrm{Var}(X) = 2 - 1^2 = 0$ and $\mathrm{Var}(Y) = 6 - 2^2 = 2$.

Hence

$$
\begin{aligned}
\mathrm{Cov}(X,Y) &= \int_0^\infty \int_0^y (x-1)(y-2)e^{-y}dxdy \\
&= \int_0^\infty \left[\left(\frac{x^2}{2}-x\right)(y-2)e^{-y}\right]_0^y dy \\
&= \frac{1}{2}\int_0^\infty (y^3-4y^2+4y)e^{-y}\ dy \\
&= \frac{1}{2}\left\{[-(y^3-4y^2+4y)e^{-y}]_0^\infty + \int_0^\infty (3y^2-8y+4)e^{-y}\ dy\right\} \\
&= \frac{1}{2}\left\{0+[-(3y^2-8y+4)e^{-y}]_0^\infty + \int_0^\infty (6y-8)e^{-y}\ dy\right\} \\
&= \frac{1}{2}\left\{0+4+[-(6y-8)e^{-y}]_0^\infty + \int_0^\infty 6e^{-y}\ dy\right\} \\
&= \frac{1}{2}\left\{0+4-8+6[-e^{-y}]_0^\infty\right\} = 1
\end{aligned}
$$

and so the correlation

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{1}{\sqrt{2}}.$$

b.

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

c. Let $U = E(X|Y)$ and $V = E(Y|X)$. Then by Part b,

$$
\begin{aligned}
U &= E(X|Y) = \frac{Y}{2} \\
V &= E(Y|X) = X+1.
\end{aligned}
$$

Let $U = g(Y) = Y/2$. Then $g^{-1}(u) = 2u$ and the density function of U is

$$f_U(u) = f_Y(g^{-1}(u))\left|\frac{d}{du}g^{-1}(u)\right| = 4ue^{-2u} \qquad (0 \le u < \infty),$$

and $f_U(u) = 0$ elsewhere. Similarly, let $h(X) = X+1$. Then $h^{-1}(v) = v-1$ and so

$$f_V(v) = f_X(h^{-1}(v))\left|\frac{d}{dv}g^{-1}(v)\right| = e^{1-v} \qquad (1 \le v < \infty),$$

and $f_V(v) = 0$ elsewhere.

Ex. 4.7.80. Let X be a continuous random variable with density function

$$f(x) = 2x, \qquad 0 \le x \le 1.$$

Find the moment-generating function of X, $M(t)$, and verify that $E(X) = M'(0)$ and that $E(X^2) = M''(0)$.

[Solution:] The moment generating function is, integrating by parts,

$$\begin{aligned}
M(t) &= \int_0^1 e^{tx} 2x \; dx \\
&= 2\left[\frac{1}{t}e^{tx}x\right]_0^1 - 2\int_0^1 \frac{1}{t}e^{tx} \; dx \\
&= \frac{2}{t}e^t - \frac{2}{t^2}\left[e^{tx}\right]_0^1 \\
&= \frac{2}{t}e^t + \frac{2}{t^2}\left(1 - e^t\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
M'(t) &= -\frac{2}{t^2}e^t + \frac{2}{t}e^t - \frac{4}{t^3}(1 - e^t) - \frac{2}{t^2}e^t \\
&= \frac{2}{t}e^t - \frac{4}{t^3}(1 - e^t) - \frac{4}{t^2}e^t \\
&= \frac{2}{t^3}(t^2e^t - 2te^t - 2 + 2e^t) \\
M''(t) &= \frac{2}{t^3}(2te^t + t^2e^t - 2e^t - 2te^t + 2e^t) - \frac{6}{t^4}(t^2e^t - 2te^t - 2 + 2e^t) \\
&= \frac{2}{t^4}(t^3e^t - 3t^2e^t + 6te^t + 6 - 6e^t)
\end{aligned}$$

Therefore, by applying L'Hospital rule once each,

$$\begin{aligned}
M'(0) &= \lim_{t\to 0} \frac{2}{t^3}(t^2e^t - 2te^t - 2 + 2e^t) \\
&= \lim_{t\to 0} \frac{2(t^2e^t + 2te^t - 2te^t - 2e^t + 2e^t)}{3t^2} \\
&= \lim_{t\to 0} \frac{2(e^t)}{3} = \frac{2}{3}, \\
M''(0) &= \lim_{t\to 0} \frac{2}{t^4}(t^3e^t - 3t^2e^t + 6te^t + 6 - 6e^t) \\
&= \lim_{t\to 0} \frac{2(3t^2e^t + t^3e^t - 6te^t - 3t^2e^t + 6e^t + 6te^t - 6e^t)}{4t^3} \\
&= \lim_{t\to 0} \frac{2t^3e^t}{4t^3} = \frac{1}{2}.
\end{aligned}$$

Evaluating directly, we find that

$$E(X) = \int_0^1 2x^2 \; dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$$

and

$$E(X^2) = \int_0^1 2x^3 \; dx = \left[\frac{2}{4}x^4\right]_0^1 = \frac{1}{2}.$$

Hence $M'(0) = E(X)$ and $M''(0) = E(X^2)$.

Ex. 4.7.85. Find the mgf of a geometric random variable, and use it to find the mean and the variance.

[Solution:]

$$M(t) \;=\; \sum_{k=1}^{\infty} e^{tk}(1-p)^{k-1}p \;=\; pe^t \sum_{k=1}^{\infty}[e^t(1-p)]^k \;=\; \frac{pe^t}{1-e^t(1-p)}.$$

Note that in the last equality, we require $|e^t(1-p)| < 1$ so that the geometric series is convergent. Hence $M(t)$ is only valid for $e^t < 1/(1-p)$, i.e., $t < \ln(1/(1-p))$.

By applying the quotient rule, we find that

$$\begin{aligned} M'(t) &= \frac{[1-e^t(1-p)]pe^t + pe^te^t(1-p))}{[1-e^t(1-p)]^2} \\ &= \frac{pe^t}{[1-e^t(1-p)]^2}, \\ M''(t) &= \frac{[1-e^t(1-p)]^2pe^t - 2pe^t[1-e^t(1-p)](-e^t(1-p))}{[1-e^t(1-p)]^4} \\ &= \frac{[1+e^t(1-p)]pe^t}{[1-e^t(1-p)]^3} \end{aligned}$$

Setting $t = 0$ gives

$$\begin{aligned} M'(0) &= \frac{p}{[1-(1-p)]^2} = \frac{1}{p}, \\ M''(0) &= \frac{[1+(1-p)]p}{[1-(1-p)]^3} = \frac{2-p}{p^2}. \end{aligned}$$

Hence

$$E(X) = \frac{1}{p} \qquad \text{and} \qquad \operatorname{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Ex. 4.7.96. If X and Y have a joint distribution, their joint moment-generating function is defined as

$$M_{XY}(s,t) = E(e^{sX+tY}),$$

which is a function of two variables, s and t.

Show how to find $E(XY)$ from the joint moment-generating function of X and Y.

[Solution:]

The partial derivatives with respect to s and then t is

$$\frac{\partial}{\partial t}\frac{\partial}{\partial s} M_{XY}(s,t) = E(XYe^{sX+tY}).$$

Setting $s = t = 0$ in the above gives $E(XY)$.