Lecture 19: Properties of Expectation

Dan Sloughter
Furman University
Mathematics 37

February 11, 2004

19.1 The unconscious statistician, revisited

The following is a generalization of the law of the unconscious statistician.

Theorem 19.1. If X and Y are discrete random variables with joint probability function p, $h : \mathbb{R} \to \mathbb{R}$, and $Y = h(X, Y)$, then

$$E[Y] = \sum_x \sum_y h(x, y)p(x, y).$$

Similarly, if X and Y are jointly continuous random variables with joint density function f, $h : \mathbb{R} \to \mathbb{R}$, and $Y = h(X, Y)$, then

$$E[Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)f(x, y)dxdy.$$

Example 19.1. Suppose X and Y have joint probability density function

$$f(x, y) = \begin{cases} 2, & \text{if } 0 < x < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then, for example,

$$E[X] = \int_0^1 \int_0^y 2xdxdy = \int_0^1 y^2 dy = \frac{1}{3}$$

and

$$E[Y] = \int_0^1 \int_0^y 2y dx dy = \int_0^1 2y^2 dy = \frac{2}{3}.$$

Note that, using the marginals f_X and f_Y which we found previously, we could have computed

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 2x(1-x) dx = 1 - \frac{2}{3} = \frac{1}{3}$$

and

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(x) dx = \int_0^1 2y^2 dy = \frac{2}{3}.$$

Example 19.2. Suppose X and Y are independent, each having a uniform distribution on $[0,1]$. Then, for example,

$$E[XY] = \int_0^1 \int_0^1 xy dx dy = \int_0^1 \frac{1}{2} y dy = \frac{1}{4}.$$

Note that $E[X] = \frac{1}{2}$ and $E[Y] = \frac{1}{2}$, so we have, in this case, $E[XY] = E[X]E[Y]$. This is in fact true in general for independent random variables.

19.2 Expectations of sums

Theorem 19.2. For any random variable X and real numbers a and b,

$$E[aX + b] = aE[X] + b.$$

Proof. We will assume X is continuous with density f (the proof for discrete X is similar). In that case,

$$E[aX+b] = \int_{-\infty}^{\infty} (ax+b) f(x) dx = a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx = aE[X] + b.$$

□

Example 19.3. Suppose X has a standard normal distribution. Then

$$E[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} x e^{-\frac{x^2}{2}}\, dx + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} x e^{-\frac{x^2}{2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \left(\lim_{b\to-\infty} \left(-e^{-\frac{x^2}{2}} \Big|_b^0 \right) + \lim_{b\to\infty} e^{-\frac{x^2}{2}} \Big|_0^b \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\lim_{b\to-\infty} \left(-1 + e^{-\frac{b^2}{2}} \right) + \lim_{b\to\infty} \left(-e^{-\frac{b^2}{2}} + 1 \right) \right) \\
&= 0.
\end{aligned}
$$

For $\sigma > 0$ and $-\infty < \mu < \infty$, let $Y = \sigma X + \mu$. Then Y is $N(\mu, \sigma^2)$, and

$$E[Y] = \sigma E[X] + \mu = \mu.$$

Theorem 19.3. If X is a random variable with moment generating function φ_X, a and b are real numbers, $Y = aX + b$, and φ_Y is the moment generating function of Y, then

$$\varphi_Y(t) = e^{tb} \varphi_X(at).$$

Proof. We have

$$\varphi_Y(t) = E[e^{t(aX+b)}] = E[e^{atX} e^{tb}] = e^{tb} \varphi_X(at).$$

□

Example 19.4. Suppose X is standard normal. Then

$$
\begin{aligned}
\varphi_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{x^2}{2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2 - 2xt)} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}((x-t)^2 - t^2)} dx \\
&= e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-t)^2} dx \\
&= e^{\frac{t^2}{2}}.
\end{aligned}
$$

Now let $Y = \sigma X + \mu$, where $\sigma > 0$ and $-\infty < \mu < \infty$. Then Y is $N(\mu, \sigma^2)$, and

$$\varphi_Y(t) = e^{\mu t} \varphi_X(\sigma t) = e^{\mu t} e^{\frac{\sigma^2 t^2}{2}} = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

Note that

$$\varphi_Y'(t) = (\mu + \sigma^2 t)e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

and

$$\varphi_Y''(t) = \left((\mu + \sigma^2 t)^2 + \sigma^2\right) e^{\mu t + \frac{\sigma^2 t^2}{2}},$$

so

$$E[Y] = \varphi_Y'(0) = \mu$$

and

$$E[Y^2] = \varphi_Y''(0) = \mu^2 + \sigma^2.$$

Theorem 19.4. For random variables X and Y and any real numbers a and b,

$$E[aX + bY] = aE[X] + bE[Y].$$

Proof. Suppose X and Y are jointly continuous with joint density f. Then

$$\begin{aligned} E[aX + bY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax + by)f(x,y)dxdy \\ &= a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy + b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dxdy \\ &= aE[X] + bE[Y]. \end{aligned}$$

□

More generally, for random variables $X_1, X_2, \ldots, X_n$ and real numbers $a_1, a_2, \ldots, a_n$, we have

$$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[x_n].$$

Example 19.5. In this example we illustrate another method for finding the expected value of a binomial random variable. First, suppose X has a Bernoulli distribution with probability of success p. Then

$$E[X] = 0 \times (1 - p) + 1 \times p = p.$$

Now suppose $X_1, X_2, \ldots, X_n$ are independent Bernoulli random variables, each with probabilty of success p, and let

$$S_n = X_1 + X_2 + \cdots + X_n.$$

Then S_n is binomial with parameters n and p. Moreover,

$$E[S_n] = E[X_1] + E[X_2] + \cdots + E[X_n] = np.$$

Example 19.6. Suppose n balls are drawn, without replacement, from an urn containing M red balls and N black balls. For $k = 1, 2, \ldots, n$, let

$$X_k = \begin{cases} 1, \text{if the } k\text{th ball is red}, \\ 0, \text{otherwise}. \end{cases}$$

Then, for any $k = 1, 2, \ldots, n$,

$$\begin{aligned} E[X_k] &= 0 \times P(X_k = 0) + 1 \times P(X_k = 1) \\ &= \frac{M(N+M-1)(N+M-2)\cdots(N+M-n+1)}{(N+M)(N+M-1)(N+M-n+1)} \\ &= \frac{M}{N+M}. \end{aligned}$$

Hence, if $S_n = X_1 + X_2 + \cdots + X_n$, then

$$E[S_n] = \frac{nM}{N+M}.$$

Note that, as in the previous example, X_1, X_2, ..., X_n are Bernoulli variables; however, in this case S_n is hypergeometric, not binomial.

19.3 Expectations of products

Theorem 19.5. If X and Y are independnent random variables, then

$$E[XY] = E[X]E[Y].$$

Proof. Suppose X and Y are jointly continuous with marginal denisties f_X and f_Y, respectively. Then

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_x(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} y f_Y(y) \int_{-\infty}^{\infty} x f_X(x) dx dy \\ &= E[X] \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= E[X]E[Y]. \end{aligned}$$

□

More generally, if X_1, X_2, $\ldots$, X_n are independent random variables, then

$$E[X_1 X_2 \cdots X_n] = E[X_1]E[X_2] \cdots E[X_n].$$

Example 19.7. If X and Y are independent $N(\mu, \sigma^2)$ random variables, then

$$E[XY] = E[X]E[Y] = \mu^2.$$

Theorem 19.6. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables, with moment generating functions φ_{X_1}, φ_{X_2}, $\ldots$, φ_{X_n}, and $Y = X_1 + X_2 + \cdots + X_n$. Then the moment generating function of Y is

$$\varphi_Y(t) = \varphi_{X_1}(t)\varphi_{X_2}(t) \cdots \varphi_{X_n}(t).$$

Proof. We have

$$\begin{aligned}
\varphi_Y(t) &= E[e^{t(X_1+X_2+\cdots+X_n)}] \\
&= E[e^{tX_1}e^{tX_2} \cdots e^{tX_n}] \\
&= E[e^{tX_1}]E[e^{tX_2}] \cdots E[e^{tX_n}] \\
&= \varphi_{X_1}(t)\varphi_{X_2}(t) \cdots \varphi_{X_n}(t).
\end{aligned}$$

□

Note that, in particular, if $X_1, X_2, \ldots, X_n$ are idependent and identically distributed (i.i.d) random variables, each with moment generating function φ, then the moment generating function of $S_n = X_1 + X_2 + \cdots + X_n$ is

$$\varphi_{S_n}(t) = (\varphi(t))^n.$$

Example 19.8. Suppose X is Bernoulli with probability of success p. The moment generating function of X is

$$\varphi_X(t) = E[e^{tX}] = (1-p) + pe^t.$$

If $X_1, X_2, \ldots, X_n$ are i.i.d. Bernoulli random variables, each with probability of success p, and $S_n = X_1 + X_n + \cdots + X_n$, then S_n is binomial and has moment generating function

$$\varphi_{S_n}(t) = (1-p+pe^t)^n.$$

19.4 Uniqueness of moment generating functions

We will find the following theorem very useful, although its proof is beyond the scope of this course.

Theorem 19.7. Suppose X and Y are random variables with moment generating functions φ_X and φ_Y, respectively. If $\varphi_X(t) = \varphi_Y(t)$ for all t in some interval $(-t_0, t_0)$, where $t_0 > 0$, then X and Y have the same distribution.

Example 19.9. Suppose X and Y are independent binomial random variables, with parameters n and p and m and p, respectively. If φ_X is the moment generating function of X, φ_Y is the moment generating function of Y, and φ_{X+Y} is the moment generating function of $X+Y$, then

$$\varphi_X(t) = (1 - p + pe^t)^n,$$

$$\varphi_Y(t) = (1 - p + pe^t)^m,$$

and

$$\varphi_{X+Y}(t) = (1 - p + pe^t)^{n+m}.$$

It follows that $X+Y$ is binomial with parameters $n+m$ and p.

Example 19.10. Suppose X and Y are independent Poisson random variables, with parameters λ and μ, respectively. If φ_X is the moment generating function of X, φ_Y is the moment generating function of Y, and φ_{X+Y} is the moment generating function of $X+Y$, then

$$\varphi_X(t) = e^{\lambda(e^t - 1)},$$

$$\varphi_Y(t) = e^{\mu(e^t - 1)},$$

and

$$\varphi_{X+Y}(t) = e^{(\lambda+\mu)(e^t - 1)}.$$

It follows that $X+Y$ is Poisson with parameter $\lambda + \mu$.

Example 19.11. Suppose X and Y are independent $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively, random variables. If φ_X is the moment generating function of X, φ_Y is the moment generating function of Y, and φ_{X+Y} is the moment generating function of $X+Y$, then

$$\varphi_X(t) = e^{\mu_X + \frac{\sigma_X^2 t^2}{2}},$$

$$\varphi_Y(t) = e^{\mu_Y + \frac{\sigma_Y^2 t^2}{2}},$$

and

$$\varphi_{X+Y}(t) = e^{\mu + \mu_Y + \frac{\left(\sigma_X^2 + \sigma_Y^2\right)t^2}{2}}.$$

It follows that $X + Y$ is $N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

Example 19.12. Suppose X and Y are independent gamma random variables, with parameters m and λ and n and λ, respectively. If φ_X is the moment generating function of X, φ_Y is the moment generating function of Y, and φ_{X+Y} is the moment generating function of $X + Y$, then

$$\varphi_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^m$$

for $t < \lambda$,

$$\varphi_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

for $t < \lambda$, and

$$\varphi_{X+Y}(t) = \left(\frac{\lambda}{\lambda - t}\right)^{m+n}$$

for $t < \lambda$. It follows that $X + Y$ is gamma with parameters $m + n$ and λ.