OCTOBER 31, 2014

LECTURE 9
EXPECTATION OF A CONTINUOUSLY DISTRIBUTED RANDOM VARIABLE, DISTRIBUTION FUNCTION AND CHANGE-OF-VARIABLE TECHNIQUES

1 Expectation of a continuously distributed random variable

Recall that in the case of a discrete random variable X distributed over the support $S_X = \{x_1, x_2, \ldots\}$ with a PMF p_X, we defined the expectation of X as a weighted average of the values in the support of the distribution with the weights given by the PMF

$$EX = \sum_{x \in S_X} x p_X(x).$$

Recall also that, in the case of continuous distributions, the role of the PMF is played by the PDF. Hence, for continuous random variables we can define the expectation similarly to that in the discrete case, however, summation must be replaced by integration (because we have continuum of values) and the PMF must be replaced by the PDF.

Definition 1. Let X be a continuously distributed random variable with the PDF $f_X(x)$. Its expected value is defined as

$$EX = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Remark. As in the case of discrete distributions, expectation is a number representing one of the properties of the distribution: its center (or location).

Example. Let $X \sim Uniform(a, b)$, $a < b$. In this case,

$$f_X(x) = \frac{1}{b-a} \times 1(a < x < b).$$

Therefore,

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_a^b x \cdot \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \cdot \left. \frac{x^2}{2} \right|_a^b \\ &= \frac{1}{b-a} \frac{b^2 - a^2}{2} \\ &= \frac{b+a}{2}. \end{aligned}$$

The result is very intuitive: since the probability mass is uniformly distributed over the interval (a, b), the center of the distribution is simply the center of the interval: $(b+a)/2 = a+(b-a)/2$.

2 Expectation of a function of a continuous random variable

2.1 The distribution function or change-of-variable technique

We often need to to evaluate expressions of the form $Eu(X)$, where X is continuously distributed and $u(x)$ is some function. Recall that, in the discrete case,

$$Eu(X) = \sum_{x \in S_X} u(x) p_X(x).$$

In the continuous case, the result is similar with necessary adjustments for continuous support:

$$Eu(X) = \int_{-\infty}^{\infty} u(x) f_X(x) dx. \tag{1}$$

The result is easy to prove in the case of monotone one-to-one functions. Recall that when a function $u(x)$ is one-to-one, it has a unique *inverse function*: if $y = u(x)$ then there is $v(y) = u^{-1}(y)$ such that $y = u(x)$ is and only if $x = v(y) = u^{-1}(y)$ or

$$v(u(x)) = u^{-1}(u(x)) = x.$$

Note that the inverse function $v(y) = u^{-1}(y)$ reverses the relationship between x and y. For example, if $y = u(x) = x^3$, then $v(y) = u^{-1}(y) = y^{1/3}$.

We will prove the following result using the definition of the CDF and by change of variable.

Theorem 2. *Let X be a continuously distributed random variable with the CDF $F_X(x)$ and the PDF $f_X(x)$. Let $Y = u(X)$, where $u(x)$ is a monotone increasing function. Then,*

(a) Y is continuously distributed with the CDF $F_X(u^{-1}(y))$ and the PDF

$$f_Y(y) = f_X(u^{-1}(y)) \frac{1}{u'(u^{-1}(y))}.$$

(b) $EY = \int_{-\infty}^{\infty} u(x) f_X(x) dx$.

Proof. We will show part (a) first. By the definition of the CDF,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(u(X) \leq y) \\ &= P(X \leq u^{-1}(y)) \\ &= F_X(u^{-1}(y)). \end{aligned}$$

Also, by the definition of the PDF,

$$\begin{aligned} f_Y(y) &= \frac{dF_Y(y)}{dy} \\ &= \frac{dF_X(u^{-1}(y))}{dy} \\ &= f_X(u^{-1}(y))\frac{1}{u'(u^{-1}(y))}, \end{aligned}$$

where $u'(x) = du(x)/dx$. For the equality in the last line we used the chain rule and the fact that[1]

$$\frac{du^{-1}(y)}{dy} = \frac{1}{u'(u^{-1}(y))}.$$

To show part (b), consider

$$EY = \int y f_Y(y) dy. \tag{2}$$

We have

$$\begin{aligned} y &= u(x), \qquad (3) \\ u^{-1}(y) &= x, \\ f_Y(y) &= f_X(u^{-1}(y))\frac{1}{u'(u^{-1}(y))} \\ &= f_X(x)\frac{1}{u'(x)}. \qquad (4) \end{aligned}$$

Lastly,

$$\begin{aligned} dy &= d(u(x)) \\ &= u'(x)dx. \qquad (5) \end{aligned}$$

Substituting (3), (4), and (5) into (2), we obtain

$$\begin{aligned} EY &= \int u(x)\left(f_X(x)\frac{1}{u'(x)}\right)\left(u'(x)dx\right) \\ &= \int u(x) f_X(x) dx. \end{aligned}$$

□

Remark. 1. The result in part (b) remains true even if $u(x)$ is not one-to-one.

[1] The result is quite intuitive: if the slope of u is u', then the slope of the inverse function at the same point is $1/u'$.

2. In part (a), if $u(x)$ is monotone decreasing, then

$$\begin{aligned} F_Y(y) &= 1 - F_X(u^{-1}(y)), \\ f_Y(y) &= f_X(u^{-1}(y)) \left| \frac{1}{u'(u^{-1}(y))} \right|. \end{aligned}$$

Example. Suppose that $X \sim Uniform(a, b)$, and let $Y = 3X$. We have $u(x) = 3x$ and $u^{-1}(y) = y/3$. Hence, by part (a) of Theorem 2,

$$\begin{aligned} f_Y(y) &= \frac{1}{b-a} \times 1(a < y/3 < b) \times \frac{1}{3} \\ &= \frac{1}{3(b-a)} \times 1(3a < y < 3b). \end{aligned}$$

We conclude that $Y \sim Uniform(3a, 3b)$, and $EY = 3(b + a)/2$.

Example. Suppose that $X \sim Uniform(a, b)$. Let

$$\begin{aligned} Y &= u(X), \text{ where} \\ u(x) &= x^3. \end{aligned}$$

In this example, $u(x)$ is monotone increasing, and

$$\begin{aligned} u^{-1}(y) &= y^{1/3} \\ u'(x) &= 3x^2, \\ u'(u^{-1}(y)) &= 3y^{2/3}, \\ f_Y(y) &= \frac{1}{b-a} \times 1(a < y^{1/3} < b) \times \frac{1}{3y^{2/3}} \\ &= \frac{1}{(b-a)3y^{2/3}} \times 1(a^3 < y < b^3). \end{aligned}$$

The distribution of Y is non-uniform: $f_Y(y)$ decreases with y over the support of the distribution (the interval (a^3, b^3)). The expected value of y is

$$\begin{aligned} EY &= \int_{a^3}^{b^3} \frac{y}{(b-a)3y^{2/3}} dy \\ &= \frac{1}{3(b-a)} \int_{a^3}^{b^3} y^{1/3} dy \\ &= \frac{1}{3(b-a)} \left. \frac{3y^{4/3}}{4} \right|_{a^3}^{b^3} \\ &= \frac{b^4 - a^4}{4(b-a)}. \end{aligned}$$

However, we can find the same expression using the result in part (b) of Theorem 2:

$$\begin{aligned} EY &= EX^3 \\ &= \int_a^b \frac{x^3}{(b-a)} dx \\ &= \frac{1}{b-a} \left. \frac{x^4}{4} \right|_a^b \\ &= \frac{b^4 - a^4}{4(b-a)}. \end{aligned}$$

2.2 Moments

Let X be a continuous random variable with the PDF f_X. Using (1), the moments of X can be defined as

$$EX^r = \int x^r f_X(x) dx.$$

Similarly,

$$\begin{aligned} Var(X) &= E(X - EX)^2 \\ &= \int (x - EX)^2 f_X(x) dx. \end{aligned}$$

Lastly, the MGF of X is given by

$$\begin{aligned} M_X(t) &= Ee^{tX} \\ &= \int e^{tx} f_X(x) dx. \end{aligned}$$

The MGF has the same properties as in the case of discrete random variables:

$$\frac{d^r M_X(t)}{dt^r} = E\left(X^r e^{tX}\right),$$

and

$$EX^r = \frac{d^r M_X(t)}{dt^r}.$$

3 Properties

In the case of continuous random variables, the same properties of expectation hold as in the case of discrete random variables. This follows from the fact that expectation remains linear whether the underlying distribution is discrete or continuous.

Theorem 3. ***(Linearity of expectation)*** *Let X be continuously distributed with the PDF*

$f_X(x)$. Let a, b_1, and b_2 be some constants. Then,

$$E(a + b_1u_1(X) + b_2u_2(X)) = a + b_1Eu_1(X) + b_2Eu_2(X).$$

Proof. By the result in (1),

$$\begin{aligned} E\left(a + b_1u_1(X) + b_2u_2(X)\right) &= \int \left(a + b_1u_1(x) + b_2u_2(x)\right) f_X(x)dx \\ &= \int af_X(x)dx + \int b_1u_1(x)f_X(x)dx + \int b_2u_2(x)f_X(x)dx \\ &= a\int f_X(x)dx + b_1\int u_1(x)f_X(x)dx + b_2\int u_2(x)f_X(x)dx \\ &= a + b_1Eu_1(X) + b_2Eu_2(X), \end{aligned}$$

where the equalities in lines 1-3 hold by the properties of integrals, and the equality in the last line holds since $\int f_X(x)dx = 1$. □

Once we have verified the linearity of expectation, all other properties based on it immediately follow. For example, consider Theorem 4 in Lecture 5 that argues that EX minimizes the function $Q(c) = E(X - c)^2$. While the result was introduced in the discussion of discrete random variables, nothing in its proof is specific to the discrete case. To prove the theorem, we only used the linearity of expectation. Thus, the theorem remains true if X is a continuous random variable:

$$EX = \arg\min_{c\in\mathbb{R}} E(X - c)^2.$$

The same applies to the properties of the variance in Theorem 2 in Lecture 6. We have not used the discreteness of a distribution in the proof and relied only on the linearity of expectation. Thus, all the properties of the variance continue to hold:

$$\begin{aligned} Var(c) &= 0, \\ Var(a + bX) &= b^2Var(X), \\ Var(X) &= EX^2 - (EX)^2. \end{aligned}$$