ESE 520 Probability and Stochastic Processes

Lecture 8

"Expectation and its properties."

Let $X : \Omega \to \mathbb{R}$ be a random variable.

Formally, the *expected value of* X (or expectation) is defined as

$$EX = \int_\Omega X(w)dP(w), \tag{1}$$

where the integral in (1) is calculated as a corresponding limit of sums of products of the values of X by the measure of a set where the value is assumed. That type of integral is called a *Lebesgue integral* and can be seen as a generalization of the concept of the usual *Riemann integral.* The Lebesgue integral is used when working with "measurable spaces" and "measurable functions". By definition, a random variable X is a measurable function acting between measurable spaces $(\Omega, \mathcal{F})$ and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

When X is a discrete random variable, then (1) becomes

$$EX = \sum_{i=1}^{\infty} x_i P(w : X(w) = x_i), \tag{2}$$

where $x_i, i = 1, 2, ...$ is the set of values of X.

In general (Ω is uncountable), one has to use the concept of integral instead of sum in (2). Luckily, there is a simple relation between a Lebesgue integral in (1) and a Riemann integral on the real line $\mathbb{R}$ (or $\mathbb{R}^n$ in general).

Recall: any random variable X generates a probability measure on Borel sets in $\mathbb{R}$ denoted by P_X which is defined as

$$P_X(B) = P(w : X(w) \in B) = P(A)$$

with $A = X^{-1}(B)$.

In particular, for $B = (-\infty, x]$, one has

$$F_X(x) = P_X((-\infty, x]) = P(A)$$

with $A = X^{-1}((-\infty, x])$.

It can be proven that

$$EX = \int_\Omega X(w)dP(w) = \int_{\mathbb{R}} x dF_X(x), \tag{3}$$

where the last integral in (3) is understood as *Riemann-Stieltjes integral*, a natural generalization of a Riemann integral on $\mathbb{R}$ when the element dx is replaced by $dF(x)$. In case when $F'_X(x) = f_X(x)$, that is when X is a continuous random variable, formula (3) becomes

$$EX = \int_\Omega X(w)dP(w) = \int_{\mathbb{R}} x f_X(x)dx \tag{4}$$

provided the last integral exists. The integral in (4) is the usual Riemann integral.

Example 1. a) Let X be a random variable that has a binomial probability distribution with parameters n and p. Then $EX = np$. Indeed:

$$EX = \sum_{i=0}^{n} i\, {}_nC_i p^i (1-p)^{n-i} = \sum_{i=1}^{n} i\, {}_nC_i p^i (1-p)^{n-i} =$$

$$np \sum_{i=1}^{n} {}_{n-1}C_{i-1} p^{i-1} (1-p)^{n-i} = np1 = np$$

since $i\, {}_nC_i = i\frac{n!}{i!(n-i)!} = n\frac{(n-1)!}{(i-1)!(n-i)!} = n\, {}_{n-1}C_{i-1}$.

b) Let X has a Poisson distribution with parameter $\lambda > 0$. Then $EX = \lambda$. Indeed:

$$EX = \sum_{i=0}^{\infty} i \frac{\lambda^i}{i!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{i=1}^{\infty} i \frac{\lambda^{i-1}}{i!} = [k = i-1] =$$

$$\lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

c) Let X has $\mathcal{N}(\mu, \sigma^2)$ distribution. Then $EX = \mu$. Indeed:

$$EX = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = [z = \frac{x-\mu}{\sigma}] =$$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}(\mu + \sigma z)e^{-\frac{1}{2}z^2} dz = \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz + 0\sigma = 1\mu = \mu.$$

In order to discuss some properties of expectation, we need first to extend the concept of functions of random variables $\phi(X)$ to the class of very general functions $\phi : \mathbb{R} \to \mathbb{R}$.

Formally, let $X : \Omega \to \mathbb{R}$ be a $\mathcal{F}$ - *random variable* and $\phi : \mathbb{R} \to \mathbb{R}$ be a map from $\mathbb{R}$ to $\mathbb{R}$. Let

$$Y := \phi(X). \tag{5}$$

What properties ϕ should have to assure that Y is again a random variable?

Definition 1. A map $\phi : \mathbb{R} \to \mathbb{R}$ is said to be a *measurable map* or a *measurable function* if

$$\phi^{-1}(B) \in \mathcal{B}(\mathbb{R})$$

for all $B \in \mathcal{B}(\mathbb{R})$. In other words, if the inverse image by ϕ of any Borel set is again a Borel set.

Using properties of the inverse image, we can write

$$Y^{-1}(B) = (\phi(X))^{-1}(B) = X^{-1}(\phi^{-1}(B)) \in \mathcal{F}$$

as long as ϕ is a measurable function.

We obtained the following: *given X is a random variable and ϕ is a measurable function, the new random variable Y defined in (5) is again a random variable on $(\Omega, \mathcal{F}, P)$.*

Remark 1. If $\phi : \mathbb{R} \to \mathbb{R}$ is a *continuous function* at a point x_0, then for any $\epsilon > 0$ there exists $\delta > 0$ such that for all x satisfying $|x - x_0| < \delta$ it holds that $|\phi(x) - \phi(x_0)| < \epsilon$. In other words, the inverse image of any open interval (that can be seen as an open interval $(\phi(x_0) - \epsilon, \phi(x_0) + \epsilon)$ around a particular point $\phi(x_0)$ is again an open interval (as an open interval around x_0). In summary: open intervals are mapped into open intervals by continuous functions.

The above means that any continuous function is a measurable function. The converse is clearly not true as a measurable function does not need to be continuous. As a simple example, we might think of an indicator function of an event A. Clearly, the measurable function $\mathbf{1}_A$ is a discontinuous one.

One can also prove that any right-continuous function (likewise left-continuous) will be measurable. Again, the converse is not necessarily true.

Some properties of expectation:

1. Let X be a continuous random variable and $\phi : \mathbb{R} \to \mathbb{R}$ is a Borel measurable function. Let $Y = \phi(X)$. Then

$$E(Y) = E[\phi(X)] = \int \phi(x) f_X(x) dx.$$

Proof: (only for the case of a "good" ϕ: ϕ has the inverse and the inverse is differentiable).

Let $f_Y(y)$ be the pdf of Y. Then we can write

$$E(Y) = \int y f_Y(y) dy = [y = \phi(x)] = \int \phi(x) f_Y(\phi(x)\phi'(x) dx =$$

$$\int \phi(x) f_X(x) dx$$

where we used the fact that X can be obtained from Y as $X = \phi^{-1}(Y)$.

2. (Linearity) For any random variables X and Y and any real numbers a and b:

$$E(aX + bY) = aE(X) + bE(Y).$$

Proof: (only for jointly continuous; for discrete: similarly) Using property 1, we obtain:

$$E(aX + bY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ax + by) f_{XY}(x, y) dx dy =$$

$$a \int_{-\infty}^{\infty} x (\int_{-\infty}^{\infty} f_{XY}(x, y) dy) dx + b \int_{-\infty}^{\infty} y (\int_{-\infty}^{\infty} f_{XY}(x, y) dx) dy =$$

$$a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} y f_Y(y) dy = aE(x) + bE(Y).$$

3. Let X and Y be independent random variables and ϕ and ψ be Borel measurable functions. Then:

$$E[\phi(X)\psi(Y)] = E[\phi(X)]E[\psi(Y)].$$

Proof: (only for jointly continuous; discrete: similarly)

$$E[\phi(X)\psi(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(x)\psi(y)f_{XY}(x,y)dxdy =$$

we use independence: $f_{XY}(x,y) = f_X(x)f_Y(y)$

$$= (\int_{-\infty}^{\infty}\phi(x)f_X(x)dx)(\int_{-\infty}^{\infty}\psi(x)f_Y(y)dy) = E[\phi(X)]E[\psi(Y)].$$

Definition 2. For any $n = 1, 2, ...,$ we define

$$\mu_n := E(X^n)$$

called *the n-th moment* of X (if it exists).

Definition 3. The number (if it exist) defined as

$$\sigma^2 = \sigma_X^2 = Var(X) := E(X - \mu_X)^2$$

is called the *variance* of X and the number $\sigma = \sqrt{\sigma^2}$ is called the *standard deviation* of X.

We can easily verify that

$$\sigma^2 = E(X^2) - \mu_1^2 = \mu_2 - \mu_1^2.$$

Example 2. Let X be a Poisson distributed random variable with parameter λ. Then $Var(X) = \lambda$.

Indeed: We can write

$$Var(X) = E(X^2) - \mu_1 = E(X^2) - \lambda^2 =$$

$$E[X(X-1)] + E(X) - \lambda^2 = E[X(X-1)] + \lambda - \lambda^2.$$

From another side,

$$E[X(X-1)] = \sum_{i=0}^{\infty} i(i-1)\frac{\lambda^i}{i!}e^{-\lambda} = \sum_{i=2}^{\infty} i(i-1)\frac{\lambda^i}{i!}e^{-\lambda} =$$

$$\lambda^2[\sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} e^{-\lambda} = \lambda^2[\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda}] = \lambda^2 1 = \lambda^2.$$

It follows then that

$$Var(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$