Probability and Stochastic Processes
Lecture 15: Expectation Revisited

Prof. Washim Uddin Mondal
Department of Electrical Engineering
Indian Institute of Technology Kanpur

Expectation in the Context of Multiple Random Variables

Example

An accident occurs at a point X that is uniformly distributed on a road of length L. At the time of the accident, an ambulance is at a location Y that is also uniformly distributed on the road. Assuming that X and Y are independent, find the expected distance between the ambulance and the point of the accident.

Solution

- The joint PDF of (X, Y) is given as:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{L^2} & \text{if } (x,y) \in (0,L)^2 \\ 0 & \text{elsewhere} \end{cases}$$

Expectation in the Context of Multiple Random Variables

Solution (Continued)

- The expected distance can be computed as:

$$
\begin{aligned}
\mathbf{E}\left[|X - Y|\right] &= \frac{1}{L^2}\int_0^L \int_0^L |x - y| \mathrm{d}x\mathrm{d}y \\
&= \frac{1}{L^2}\int_0^L \int_0^y (y - x)\mathrm{d}x\mathrm{d}y + \frac{1}{L^2}\int_0^L \int_y^L (x - y)\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2L^2}\int_0^L y^2 \mathrm{d}y + \frac{1}{2L^2}\int_0^L (L - y)^2 \mathrm{d}y = \frac{L}{3}
\end{aligned}
$$

Expectation in the Context of Multiple Random Variables

Example

Find (a) $\mathbf{E}[XY]$ and (b) $\mathbf{E}[X]$ if the joint PDF of (X, Y) be given as:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{y} & \text{if } 0 < x < y, 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Solution

$$\mathbf{E}[XY] = \int_0^1 \int_0^y xy \times \frac{1}{y} \mathrm{d}x\mathrm{d}y = \frac{1}{6}, \quad \mathbf{E}[X] = \int_0^1 \int_0^y \frac{x}{y} \mathrm{d}x\mathrm{d}y = \frac{1}{4}$$

Expectation of Sum of Random Variables

Proposition 15.1

Let $X_1, \cdots, X_n$ be arbitrary random variables defined on the same sample space $(\Omega, \mathcal{F}, \mathbf{P})$. The following relation must hold true.

$$\mathbf{E}[X_1 + \cdots + X_n] = \mathbf{E}[X_1] + \cdots + \mathbf{E}[X_n]$$

Proof of Proposition 15.1 (Special Case)

- Assume that f_{X_1,X_2} is the joint PDF of (X_1, X_2).

$$\mathbf{E}[X_1 + X_2] = \int_{\mathbb{R}^2} (x_1 + x_2) f_{X_1,X_2}(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

Expectation of Sum of Random Variables

Proof of Proposition 15.1 (Special Case)

$$
\begin{aligned}
\mathbf{E}[X_1 + X_2] &= \int_{\mathbb{R}} x_1 \int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\mathrm{d}x_2\mathrm{d}x_1 + \int_{\mathbb{R}} x_2 \int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 \\
&= \int_{\mathbb{R}} x_1 f_{X_1}(x_1)\mathrm{d}x_1 + \int_{\mathbb{R}} x_2 f_{X_2}(x_2)\mathrm{d}x_2 = \mathbf{E}[X_1] + \mathbf{E}[X_2]
\end{aligned}
$$

- A similar technique can be used for discrete random variables.
- Result can be extended to n random variables.
- General proof uses the properties of Lebesgue integration.

Expectation of Sum of Random Variables

Example

Let X be a binomial random variable with parameters (n, p). Recall that $X = X_1 + \cdots + X_n$ where $X_1, \cdots, X_n$ are independent Bernoulli random variables with the same parameter p. Therefore, $\mathbf{E}[X] = \mathbf{E}[X_1] + \cdots + \mathbf{E}[X_n] = np$.

Example

Let X be a negative binomial random variable with parameters (r, p). Recall that $X = X_1 + \cdots + X_r$, where $X_1, \cdots, X_r$ are independent geometric random variables with the same parameter p. Therefore, $\mathbf{E}[X] = \mathbf{E}[X_1] + \cdots + \mathbf{E}[X_r] = \frac{r}{p}$.

Expectation of Sum of Random Variables

Example

N people throw their hats into the center of a room. The hats are mixed up, and each person randomly selects one. Find the expected number of people who select their own hat.

Solution

- Let X be the desired number. We get $X = X_1 + \cdots + X_N$, where

$$X_i = \begin{cases} 1 & \text{if the } i\text{th person selects his/her own hat} \\ 0 & \text{elsewhere} \end{cases}$$

- Note that $\mathbf{E}[X_i] = \mathbf{P}(X_i = 1) = (N-1)!/N! = 1/N$. Thus, $\mathbf{E}[X] = 1$.

Covariance and Correlation

Definition

Let X, Y be arbitrary random variables. Their covariance is defined as:

$$\mathrm{Cov}(X, Y) = \mathbf{E}\left[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])\right] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = \mathrm{Cov}(Y, X)$$

If $\mathrm{Var}(X) > 0$ and $\mathrm{Var}(Y) > 0$, then the correlation coefficient of (X, Y) is:

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

Note

- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$ for any random variable X.

Covariance and Correlation

Notes

- $\rho(X,X) = 1$ for any random variable X with $\mathrm{Var}(X) > 0$.
- $\mathrm{Cov}(a_1X_1 + b_1, a_2X_2 + b_2) = a_1a_2\mathrm{Cov}(X_1, X_2)$.
- If $a_1, a_2 \neq 0$, then $\rho(a_1X_1 + b_1, a_2X_2 + b_2) = \mathrm{sgn}(a_1a_2)\rho(X_1, X_2)$.
- If $a_1, a_2 \neq 0$, and $\mathrm{Var}(X) > 0$, then $\rho(a_1X + b_1, a_2X + b_2) = \mathrm{sgn}(a_1a_2)$.

Proposition 15.2

$$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}\mathrm{Cov}(X_i, Y_j)$$

Covariance and Correlation

Proof of Proposition 15.2

$$
\begin{aligned}
\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) &= \mathbf{E}\left[\left\{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathbf{E}[X_i]\right\}\left\{\sum_{j=1}^{m} Y_j - \sum_{j=1}^{m} \mathbf{E}[Y_j]\right\}\right] \\
&= \mathbf{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{m}(X_i - \mathbf{E}[X_i])(Y_j - \mathbf{E}[Y_j])\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \mathbf{E}\left[(X_i - \mathbf{E}[X_i])(Y_j - \mathbf{E}[Y_j])\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}(X_i, Y_j)
\end{aligned}
$$

Covariance and Correlation

Corollary 15.3

If $X_1, \cdots, X_n$ are arbitrary random variables, then

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)$$

Proof of Corollary 15.3

$$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n} \mathrm{Cov}(X_i, X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)$$

Covariance and Correlation

Proposition 15.4

If X, Y are independent random variables and $g, h : \mathbb{R} \to \mathbb{R}$ are arbitrary Borel-measurable functions, then $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)]\mathbf{E}[h(Y)]$.

Proof of Proposition 15.4 (Continuous Case)

- Let the joint PDF of (X, Y) be $f_{X,Y}$. We have $f_{X,Y}(x, y) = f_X(x)f_Y(y), \ \forall (x, y) \in \mathbb{R}^2$.

$$
\begin{aligned}
\mathbf{E}\left[g(X)h(Y)\right] &= \int_{\mathbb{R}^2} g(x)h(y)f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y \\
&= \left\{\int_{\mathbb{R}} g(x)f_X(x)\mathrm{d}x\right\}\left\{\int_{\mathbb{R}} h(y)f_Y(y)\mathrm{d}y\right\} = \mathbf{E}[g(X)]\mathbf{E}[h(Y)]
\end{aligned}
$$

Covariance and Correlation

Proposition 15.5

If X, Y are independent random variables, then $\mathrm{Cov}(X, Y) = 0$. Moreover, if $\mathrm{Var}(X) > 0$ and $\mathrm{Var}(Y) > 0$, then $\rho(X, Y) = 0$.

Proof of Proposition 15.5

$$\mathrm{Cov}(X, Y) = \mathbf{E}\left[(X - \mathbf{E}[X])(Y - \mathrm{E}[Y])\right] = \mathbf{E}[X - \mathbf{E}[X]]\mathbf{E}[Y - \mathbf{E}[Y]] = 0$$

Notes

- The converse is false, i.e., zero correlation $\not\Rightarrow$ independence.
- If X_i's are independent, then $\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$.

Covariance and Correlation

Example

Find $\text{Cov}(X, X^2)$, where $X \sim \mathcal{N}(0, 1)$.

Solution

- Note that $\text{Cov}(X, X^2) = \mathbf{E}[X^3] - \mathbf{E}[X]\mathbf{E}[X^2]$.
- Recall that $\mathbf{E}[X] = 0$ and $\mathbf{E}[X^3] = 0$. Hence $\text{Cov}(X, X^2) = 0$.

Note

- Clearly, X and X^2 are not independent, though $\text{Cov}(X, X^2) = 0$.

Covariance and Correlation

Example

Let $X_1, \cdots, X_n$ be independent and identically distributed (IID) random variables with the same mean μ and the same variance σ^2. The sample mean, $\bar{X}_n$, and the sample variance, S_n^2, of such a collection are:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

Find (a) $\mathbf{E}[\bar{X}_n]$, (b) $\mathrm{Var}(\bar{X}_n)$, and (c) $\mathbf{E}[S_n^2]$.

Solution

- $\mathbf{E}[\bar{X}_n] = \sum_{i=1}^{n} \mathbf{E}[X_i]/n = \mu.$

Solution (Continued)

- The variance of $\bar{X}_n$ can be computed as:

$$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) + \frac{2}{n^2}\sum_{1\le i<j\le n}\mathrm{Cov}(X_i, X_j) = \frac{\sigma^2}{n}$$

- Note that $X_i - \bar{X}_n = (X_i - \mu) + (\mu - \bar{X}_n)$. Hence,

$$\mathbf{E}[(X_i - \bar{X}_n)^2] = \mathrm{Var}(X_i) + 2\mathbf{E}[(X_i - \mu)(\mu - \bar{X}_n)] + \mathrm{Var}(\bar{X}_n)$$

- Recall that $\mathrm{Var}(X_i) = \sigma^2$ and $\mathrm{Var}(\bar{X}_n) = \sigma^2/n$.

Covariance and Correlation

Solution (Continued)

$$\mathbf{E}\left[(X_i-\mu)(\mu-\bar{X}_n)\right] = \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}\left[(X_i-\mu)(\mu-X_j)\right]$$

$$= -\frac{1}{n}\mathrm{Var}(X_i) - \frac{1}{n}\sum_{j\neq i}\mathrm{Cov}(X_i,X_j) = -\frac{\sigma^2}{n}$$

- In summary, we have the following:

$$\mathbf{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\right] = \frac{n}{n-1}\left(\sigma^2-\frac{2\sigma^2}{n}+\frac{\sigma^2}{n}\right) = \sigma^2$$

Covariance and Correlation

Proposition 15.6

If $\mathrm{Var}(X) > 0$ and $\mathrm{Var}(Y) > 0$ for any two random variables X, Y, then $-1 \leq \rho(X, Y) \leq 1$.

Proof of Proposition 15.6

- Note that $\mathrm{Var}(aX + Y) = a^2\mathrm{Var}(X) + 2a\mathrm{Cov}(X, Y) + \mathrm{Var}(Y)$.
- Moreover, $\mathrm{Var}(aX + Y) \geq 0, \forall a \in \mathbb{R}$.
- Therefore, the discriminant obeys $4\left[\mathrm{Cov}(X, Y)\right]^2 - 4\mathrm{Var}(X)\mathrm{Var}(Y) \leq 0$.
- This leads to $|\mathrm{Cov}(X, Y)| \leq \sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$.
- Hence, $|\rho(X, Y)| \leq 1$.

Covariance and Correlation

Example

Let X be a binomial random variable with parameters (n, p). Note that $X = X_1 + \cdots + X_n$, where X_i's are independent Bernoulli random variables with parameter p. Hence, we have: $\text{Var}(X) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) = np(1-p)$.

Example

Let X be a negative binomial random variable with parameters (r, p). Note that $X = \sum_{i=1}^{r} X_i$, where X_i's are independent geometric random variables with the same parameter p. We have the following: $\text{Var}(X) = \text{Var}(X_1) + \cdots + \text{Var}(X_r) = r(1-p)/p^2$.

Moment Generating Function (MGF)

Proposition 15.7

If X, Y are independent random variables with M_X and M_Y as their respective MGFs, then the MGF of $X+Y$ is given as follows $\forall t \in \mathbb{R}$.

$$M_{X+Y}(t) = \mathbf{E}[e^{t(X+Y)}] = \mathbf{E}[e^{tX}]\mathbf{E}[e^{tY}] = M_X(t)M_Y(t)$$

Example

If $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$ are independent, then

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = \exp((\lambda_1+\lambda_2)(e^t-1))$$

for all $t \in \mathbb{R}$. Hence, $X_1 + X_2 \sim \text{Poisson}(\lambda_1+\lambda_2)$.

Moment Generating Function (MGF)

Example

If $X_1 \sim \text{Binomial}(n_1, p)$, and $X_2 \sim \text{Binomial}(n_2, p)$ are independent, then

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = (1 - p + pe^t)^{n_1+n_2}$$

for all $t \in \mathbb{R}$. Hence, $X_1 + X_2 \sim \text{Binomial}(n_1 + n_2, p)$.

Example

If $X_1 \sim \text{NB}(r_1, p)$ and $X_2 \sim \text{NB}(r_2, p)$ are independent, then

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = (pe^t/(1 - e^t(1 - p)))^{r_1+r_2}$$

for all $t < -\ln(1 - p)$. Hence, $X_1 + X_2 \sim \text{NB}(r_1 + r_2, p)$.

Moment Generating Function (MGF)

Example

If $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent, then

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = \exp\left((\mu_1 + \mu_2)t + \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}\right)$$

for all $t \in \mathbb{R}$. Hence, $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Example

If $X_1 \sim \text{Gamma}(\alpha_1, \lambda)$ and $X_2 \sim \text{Gamma}(\alpha_2, \lambda)$ are independent, then

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha_1-\alpha_2}$$

for all $t < \lambda$. Hence, $X_1 + X_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \lambda)$.