Chapter 4 - Random Variables and Distributions

STAT 251

Lecture 10

Properties of Mean and Varience and Covariance,

Sum of Independent Random Variables,

and

Maximum & Minimum of Independent Random Variables

Dr. Lasantha Premarathna

Chapter 4 - Learning Outcomes

- Notations
- Discrete Random Variables
- Continuous Random Variables
- Probability mass function (pmf)
- Probability density function (pdf)
- Cumulative Distribution Function (cdf)
- The Mean, the Variance and the Standard Deviation, Covariance
- Max and Min of Independent Random Variables

Properties of the Mean and Variance

1. $E(aX + b) = aE(X) + b$ for all constants a and b
2. $E(X + Y) = E(X) + E(Y)$ for all pairs of random variables X and Y.
3. $E(XY) = E(X)E(Y)$ for all pairs of **independent random variables** X and Y $\boxed{i}$

   → only shift.
4. $Var(aX + b) = a^2 Var(X)$ for all constants a and b
5. If X and Y are **independent random variables** $\boxed{i}$

$$Var(X + Y) = Var(X) + Var(Y)$$
$$Var(X - Y) = Var(X) + Var(Y)$$

↳ $= Var(x + (-1)y)$
$= Var(x) + (-1)^2 Var(y)$
$= Var(x) + Var(y)$

Var still incr. w/ & var.

Proof: $Var(X) = E(X^2) - [E(X)]^2$

$$\begin{aligned}
Var(X) &= E[(X - E(X))^2] \\
&= E[(X - \mu)^2] \\
&= E[X^2 - 2X\mu + \mu^2] \\
&= E[X^2] - E[2X\mu] + E[\mu^2] \\
&= E[X^2] - 2\mu E[X] + E[\mu^2] \qquad ; E[X] = \mu \\
&= E[X^2] - 2\mu^2 + \mu^2 \quad \text{✓ const. value.} \\
&= E[X^2] - \mu^2 \\
&= E(X^2) - [E(X)]^2
\end{aligned}$$

Covariance

- In bivariate setting involving random variavles X and Y, covariance is given by

$$Cov(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$$

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

- Covariance is positive when large X's tend to occur with large Y's and when small X's tend to occur with small Y's. Similarly covariance is negative when large X's tend to occur with small Y's and when small X's tend to occur with large Y's.

- If X and Y are **independent random variables**

$$Cov(X, Y) = 0$$

Covariance - for non-independent

- $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
- More generally

  Const does not matter.

  $$Var(aX+bY+c) = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$$
- if X and Y are **independent random variables**

  $$Var(aX+bY+c) = a^2Var(X) + b^2Var(Y)$$

Example 8

Consider two random variables X and Y. Given that $Var(X) = 5$, $Var(Y) = 10$ and $Cov(X, Y) = 2$. Find the variance of W such that

$$W = 2X + 3Y$$

$$Var(2x + 3y) = 4\,Var(x) + 9\,Var(y) + 2(2)(3)\,Cov(x, y)$$

$$= 20 + 90 + 24$$

$$= 134 \,//$$

Example 9

Consider two random variables X and Y. Given that $Var(X) = 5$, $Var(Y) = 10$ and $Cov(X, Y) = 2$. Find the variance of U such that

$$U = 2X - Y$$

$$Var(2x - y) = 4\,Var(x) + Var(y) + 2(2)(-1)\,Cov(x, y)$$

$$= 4 \times 5 + 10 - 4 \times 2$$

$$20 + 10 - 8$$

$$= 22 //$$

Sum of Independent Random Variables

Random Experiments are often independently repeated many times generating a sequence $X_1, X_2, \cdots, X_n$ of n independent random variables. We will consider a linear combination of these variables

$$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

where $a_1, a_2, \cdots, a_n$ are constants

Then

$$E(Y) = a_1 E(X_1) + a_2 E(X_2) + \cdots + a_n E(X_n)$$

$$Var(Y) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + \cdots + a_n^2 Var(X_n)$$

Sum of Independent Random Variables (Contd...)

If the n random variable X_i have common mean μ and common variance σ^2, we get

$$E(Y) = (a_1 + a_2 + \cdots + a_n)\,\mu$$

$$Var(Y) = (a_1^2 + a_2^2 + \cdots + a_n^2)\,\sigma^2$$

In this case, the sequence $X_1, X_2, \cdots, X_n$ ia said to be a random sample.

Average of Independent Random Variables

$X_1, X_2, \cdots, X_n$ are n independent variables and

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Then

$$E(\bar{X}) = \frac{1}{n}\{E(X_1) + E(X_2) + \cdots + E(X_n)\}$$

$$Var(\bar{X}) = \frac{1}{n^2}\{Var(X_1) + Var(X_2) + \cdots + Var(X_n)\}$$

In the n random variables X_i have a common mean μ and common variance σ^2, then

$$E(\bar{X}) = \mu \qquad Var(\bar{X}) = \frac{\sigma^2}{n}$$

Average of Independent Random Variables

$$E(\bar{X}) = \mu \qquad Var(\bar{X}) = \frac{\sigma^2}{n}$$

$$\begin{aligned}
E(\bar{X}) &= E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n}E\left(X_1 + X_2 + \cdots + X_n\right) \\
&= \frac{1}{n}\left(E(X_1) + E(X_2) + \cdots + E(X_n)\right) = \frac{1}{n}\left(\mu + \mu + \cdots + \mu\right) \\
&= \frac{1}{n}\, n\mu = \mu
\end{aligned}$$

$$\begin{aligned}
Var(\bar{X}) &= Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n^2}Var\left(X_1 + X_2 + \cdots + X_n\right) \\
&= \frac{1}{n^2}\left(Var(X_1) + Var(X_2) + \cdots + Var(X_n)\right) \\
&= \frac{1}{n^2}\left(\sigma^2 + \sigma^2 + \cdots + \sigma^2\right) = \frac{1}{n^2}\, n\sigma^2 = \frac{\sigma^2}{n}
\end{aligned}$$

Maximum of Independent Random Variables

The maximum of a sequence of n independent random variables are of practical interest

The **maximum** $V = max\{X_1, X_2, \cdots, X_n\}$ can be used to model

- The lifetime of a system on n components connected in Parallel

  X_i = lifetime of the i^{th} component

- The maximum flood level of a river in the next n years

  X_i = maximum flood level in the i^{th} year

Maximum of Independent Random Variables

Given that n independent random variables are $X_1, X_2, \cdots, X_n$. X_i's are identically distributed with the pdf $f_X(x)$ and cdf $F_X(x)$.

Find the pdf of the maximum

$$V = max\{X_1, X_2, \cdots, X_n\}$$

first, find the cdf of V, $F_V(v)$

$$\begin{aligned} F_V(v) &= P(V \le v) \\ &= P(X_1 \le v, X_2 \le v, \cdots, X_n \le v) \quad \text{; if } v \text{ is the maximum,} \\ &\qquad\qquad\qquad\qquad \text{then each of } X_1, X_2, \cdots, X_n \text{ are } \le v \\ &= P(X_1 \le v)P(X_2 \le v) \cdots P(X_n \le v) \quad ; X_i\text{'s are independent} \\ &= F_{X_1}(v)F_{X_2}(v) \cdots F_{X_n}(v) \\ F_V(v) &= [F_X(v)]^n \quad ; X_i\text{'s are identically distributed} \\ &\qquad \text{therefore } F_{X_1}(v) = F_{X_2}(v) = \cdots = F_{X_n}(v) = F_X(v) \end{aligned}$$

Maximum of Independent Random Variables

$$V = max\{X_1, X_2, \cdots, X_n\}$$

pdf of V is $f_V(v)$

$$\begin{aligned}
f_V(v) &= F_V^{'}(v) \\
&= \frac{d}{dv} F_V(v) \\
&= \frac{d}{dv} \left[F_X(v)\right]^n \\
&= n \left[F_X(v)\right]^{n-1} \frac{d}{dv} F_X(v) \\
f_V(v) &= n \left[F_X(v)\right]^{n-1} f_X(v)
\end{aligned}$$

Minimum of Independent Random Variables

The minimum of a sequence of n independent random variables are of practical interest

The **minimum** $U = min\{X_1, X_2, \cdots, X_n\}$ can be used to model

- The lifetime of a system on n components connected in series
  X_i = lifetime of the i^{th} component
- The minimum flood level of a river in the next n years
  X_i = minimum flood level in the i^{th} year

Minimum of Independent Random Variables

Given that n independent random variables are $X_1, X_2, \cdots, X_n$. X_i's are identically distributed with the pdf $f_X(x)$ and cdf $F_X(x)$.
Find the pdf of the minimum

$$U = min\{X_1, X_2, \cdots, X_n\}$$

first, find the cdf of U, $F_U(u)$

$$\begin{aligned}
F_U(u) &= P(U \leq u) \\
&= 1 - P(U > u) \\
&= 1 - P(X_1 > u, X_2 > u, \cdots, X_n > u) \quad ;\text{if } u \text{ is the minimum,} \\
&\qquad\qquad\qquad\qquad\qquad \text{then each of } X_1, X_2, \cdots, X_n \text{ are } > u \\
&= 1 - P(X_1 > u)P(X_2 > u) \cdots P(X_n > u) \;\; ; X_i\text{'s are independent} \\
&= 1 - [1 - F_{X_1}(u)][1 - F_{X_2}(u)] \cdots [1 - F_{X_n}(u)] \\
F_U(u) &= 1 - [1 - F_X(u)]^n \quad ; X_i\text{'s are identically distributed}
\end{aligned}$$

Minimum of Independent Random Variables

$$U = min\{X_1, X_2, \cdots, X_n\}$$

pdf of U is $f_U(u)$

$$\begin{aligned}
f_U(u) &= F_U^{'}(u) \\
&= \frac{d}{du} F_U(u) \\
&= \frac{d}{du}\left\{1 - [1 - F_X(u)]^n\right\} \\
&= 0 - n\left[1 - F_X(u)\right]^{n-1} \frac{d}{du}(-F_X(u))
\end{aligned}$$

$$f_U(u) = n\left[1 - F_X(u)\right]^{n-1} f_X(u) \qquad ; \frac{d}{du}(-F_X(u)) = -\frac{d}{du}F_X(u) = -f_X(u)$$

Before the next class ...

Visit the course website at canvas.ubc.ca

- Review the lecture 10 and related sections in the text book
- Topic of next class: **More on continuous rvs, Max & Min of independent rvs, Examples, and in-class Activity**