5.8 Linear Functions of Random Variables

- In statistical studies, we often use statistics that are linear functions of random measurements $Y_1, Y_2, \ldots, Y_n$. That is, for some constants $a_1, \ldots, a_n$:

Theorem: Consider r.v.'s $Y_1, \ldots, Y_n$ and $X_1, \ldots, X_m$ such that $E(Y_i^2) < \infty$ and $E(X_j^2) < \infty$ for all i, j. For constants $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$, define:

$$U_1 = \sum_{i=1}^{n} a_i Y_i \quad \text{and} \quad U_2 = \sum_{j=1}^{m} b_j X_j$$

Then:

(1) $$E(U_1) = \sum_{i=1}^{n} a_i E(Y_i) = a_1 E(Y_1) + \cdots + a_n E(Y_n)$$

(2) $$V(U_1) = \sum_{i=1}^{n} a_i^2 V(Y_i) + 2 \sum_{i<j}\sum a_i a_j \operatorname{cov}(Y_i, Y_j)$$

$$= \sum_{i=1}^{n} a_i^2 V(Y_i) + \sum_{i \neq j}\sum a_i a_j \operatorname{cov}(Y_i, Y_j)$$

(3) $$\operatorname{cov}(U_1, U_2) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \operatorname{cov}(Y_i, X_j)$$

Special Cases of this Theorem:

(i)

(ii)

(iii) For constants a_1 and a_2,

Proof of (iii): [Note (i) and (ii) are simply specific cases of (iii).]

(iv) For constants a_1, a_2, b_1, b_2,
$\text{cov}(a_1 Y_1 + a_2 Y_2, b_1 X_1 + b_2 X_2) =$

Corollary: If Y_1 and Y_2 are independent r.v.'s:
$V(Y_1 + Y_2) =$
$V(Y_1 - Y_2) =$
$V(a_1 Y_1 + a_2 Y_2) =$
- This result follows immediately from (i)-(iii), since if Y_1, Y_2 independent, then

Example 7 again: Find $V(Y_2 - Y_1)$.
- We have shown that

So

Find $\text{cov}(Y_1, Y_2 - Y_1)$ and $\text{cov}(Y_1 + Y_2, Y_2 - Y_1)$.

5.11 Conditional Expectations

Defn. For two r.v.'s Y_1 and Y_2, the conditional expectation of $g(Y_1)$, given that $Y_2 = y_2$, is:

for jointly continuous Y_1, Y_2,

and

for jointly discrete Y_1, Y_2.

Example 6 again: Find $E(Y_2 \mid Y_1 = 0.7)$ and $V(Y_2 \mid Y_1 = 0.7)$.

- We have found $f(y_2 \mid y_1) = \dfrac{y_1 + 4y_2}{y_1 + 2}$ for $0 < y_2 < 1$.

So

-We can also consider conditional expectation $E(Y_1|Y_2)$ and conditional variance $V(Y_1|Y_2)$ letting Y_2 range over its possible values, rather than fixing Y_2 at a specific value.

-While $E(Y_1|Y_2=y_2)$ is a constant, $E(Y_1|Y_2)$ is a random variable, since it depends on Y_2.

<u>Theorem</u> (Law of Iterated Expectation): If Y_1 and Y_2 are any r.v.'s, then:

-Note that the inside expected value is with respect to the conditional distribution of $Y_1|Y_2$ and the outside expected value is with respect to the distribution of Y_2.

Proof (Continuous case):

- The proof is similar in the discrete case.

Theorem (Law of Iterated Variance):
If Y_1 and Y_2 are any r.v.'s, then

Proof:

Example: Let Y be the number of defective parts from an assembly line out of 20 produced per day. Suppose the probability p of a part being defective depends on the day and follows a beta$(1, 9)$ distribution. Find $E(Y)$ and $V(Y)$.