More Expectations and Variances (P.429~P.455)

Let X *have probability mass function p, and* Y *have probability density function f, then the expectation could be computed as*

$E(X) = \sum_{x \in A} xp(x),\ E(Y) = \int_{-\infty}^{\infty} yf(y)dy,$

(Theorem 10.1, P.429) For random variables $X_1, X_2, \cdots, X_n$ defined on the same sample space, and for real numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$,

$E(\sum_{i=1}^{n} \alpha_i X_i) = \sum_{i=1}^{n} \alpha_i E(X_i)$

Proof for continuous case.

$$E\left(\sum_{i=1}^{n} \alpha_i X_i\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\sum_{i=1}^{n} \alpha_i x_i\right) \mathrm{f}(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n$$

$$= \sum_{i=1}^{n} \alpha_i \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n = \sum_{i=1}^{n} \alpha_i E(X_i)$$

Corollary and Examples 10.1, 10.2 on P.430

(Corollary) Let $X_1, X_2, \cdots, X_n$ be random variables on the same sample space. Then

$$E(X_1 + X_2 + \cdots X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

(10.1) A die is rolled 15 times. What is the expected value of the sum of the outcomes? Ans: E(X_i)=$\sum_{x=1}^{6} xp(x) = \frac{7}{2}$, thus, E(X)=15 $\times$ ($\frac{7}{2}$)=52.5.

(10.2) A well-shuffled ordinary deck of 52 cards is divided randomly into four piles of 13 each. Counting jack, queen, and king as 11, 12, and 13, respectively. We say that a match occurs in a pile if the *j*th card is *j*.

What is the expected value of the total number of matches in all four piles? Ans: E(X)=4 on P.431.

Example 10.3 on P.431 (Daniel Bernoulli 1700-1782)

(10.3) Exactly *n* married couples are living in a small town. What is the expected number of intact (完整的) couples after *m* deaths among the *n* couples? Assume that the deaths occur at random, there are no divorces, and there are no new marriages.

Ans: Let

Let X $be\ the\ number\ of\ intact\ couples\ after\ m\ deaths, and\ for\ i = 1, 2, \cdots, n\ define$

$$X_i = \begin{cases} 1 & if\ the\ ith\ couple\ is\ left\ intact \\ 0 & otherwise \end{cases}$$

Then $X = \sum_{i=1}^{n} X_i$, and $E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1)$,

Thus, $\mathrm{E(X)} = nP(X_i = 1) = \mathrm{n} \times \binom{2n-2}{m} / \binom{2n}{m}$ since the ith couple remain.

Examples 10.6 and 10.7 on P.432-433.

(10.6) Let X be a binomial random variable with parameters (n, p). Recall that X is the number of successes in n independent Bernoulli trials. Thus, let $X_i = 1$ be the ith trial is success and $X_i = 0\ otherwise$.

Then $X = X_1 + X_2 + \cdots + X_n, and\ E(X) = np$.

(10.7) Let X be a negative binomial random variable with parameters (r, p). Then in a sequence of independent Bernoulli trials each with success probability p, X is the number of trials until the rth success. Let X_1 be the number of trials until the first success, X_2 be the number of additional trials to get the 2nd success, and so on. Then

$X = X_1 + X_2 + \cdots + X_r$, where for $i = 1,2, \cdots, n$, the random variable X_i is geometric with parameter $p\ because\ P(X_i = n) = (1-p)^{n-1}p$

by the independence of the trials. Therefore, E(X)=$\frac{r}{p}$

~~Theorem 10.2 on P.434~~ (Like a paradox)

If $\sum_{i=1}^{\infty} E(|X_i|) < \infty$ or if, for all i, the random variables $X_1, X_2, \cdots$ are nonnegative, (that is, $P(X_i \geq 0) = 1\ for\ i \geq 1$), then

$E(\sum_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} E(X_i)$

(Theorem 10.2) Let *N* be a discrete random variable with set of possible values {1, 2, 3, ⋯ , }. Then

$$E(N) = \sum_{i=1}^{\infty} P(N \geq i).$$

$E(N) = \sum_{i=1}^{\infty} P(N \geq i)$. on P.434

- $For\ i \geq 1, let$
- $X_i = \begin{cases} 1 & if\ N \geq i \\ 0 & Otherwise; \end{cases}$
- Then $\sum_{i=1}^{\infty} X_i = \sum_{i=1}^{N} X_i + \sum_{i=N+1}^{\infty} X_i = \sum_{i=1}^{N} 1 + \sum_{i=N+1}^{\infty} 0 = N.$
- Since $P(X_i \geq 0) = 1, for\ i \geq 1,$
- $E(N) = E(\sum_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} E(X_i) = \sum_{i=1}^{\infty} P(N \geq i).$

Note that if X is continuous nonnegative random variable, then

$E(X) = \int_0^{\infty} P(X > t)dt$ (See Remark 6.4 on P.262)

(Theorem 10.3) Cauchy-Schwarz Inequality on P.435

(Theorem 10.3) For random variables X *and* Y with finite variances,

$$|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$$

Proof: For any $t > 0, (X - tY)^2 \geq 0$. Hence for all values of t,
$X^2 - 2XYt + t^2Y^2 \geq 0$, and $\mathrm{g(t)} = \mathrm{E}(X^2 - 2XYt + t^2Y^2) \geq 0$, for any $t, including\ t = 0$, which implies

$$E(Y^2)t^2 - 2E(XY)t + E(X^2) \geq 0$$

(Since $g'(t) = 0\ implies\ that\ t = E(XY)/E(Y^2)$, plug-in to get it).

(Corollary) For a random variable X, $[E(X)]^2 \leq E(X^2)$.

10.2 Covariance on P.439

- $E(aX + bY) = aE(X) + bE(Y)$; $Var(X) = E[(X - E(X))^2] = E(X^2) - (E(X))^2$
- $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abE[(X - E(X))(Y - E(Y))]$

(Definition 10.1, P.439) *Let X, Y be jointly distributed random variables; then the covariance of X and Y is defined by*

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

In particular, Cov(X,X)=Var(X), and |Cov(X,Y)| $\leq \sigma_X \sigma_Y$, Cov(X,Y)= $\rho_{XY}\sigma_X\sigma_Y$

(Theorem 10.4) Let *a, b*∈ R; for random variables X and Y, we have

$$Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y).$$

(Proof) $Var(aX + bY) = E\left[\left((aX + bY) - E(aX + bY)\right)^2\right]$

$$= E[(a\left(X - E(X)\right) + b(Y - E(Y)))^2]$$

Equation 10.8 on P.441

- $Cov(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j Cov(X_i, Y_j)$
- For random variables X and Y, Cov(X,Y) might be positive, negative, or zero. If Cov(X,Y)=0, we say that X and Y are *uncorrelated*. If X and Y are independent, then X and Y are uncorrelated, but the converse may not be true.

(10.9) Let X be uniformly distributed over (-1, 1), that is f(x)=1/2 when x in (-1,1), then $E(X) = 0$, and $E(X^3) = 0$. Define $Y = X^2$. Then $Cov(X,Y) = E(X^3) - E(X)E(X^2) = 0$, thus, X and Y are uncorrelated but are not independent.

Examples 10.10 on P.441-P.442

(10.10) There are 300 cards in a box numbered 1 through 300. Therefore, the number on each card has one, two, or three digits. A card is drawn randomly from the box. Suppose that the number on the card has X digits of which Y are 0. Determine whether X and Y are positively correlated, negatively correlated, or uncorrelated. Ans: First, find $p(x,y), x = 1,2,3; y = 0,1,2,$ then

E(X)= $1 \cdot \frac{9}{300} + 2 \cdot \frac{90}{300} + 3 \cdot \frac{201}{300} =$ 2.64,

E(Y)=$0 \cdot \frac{252}{300} + 1 \cdot \frac{45}{300} + 2 \cdot \frac{3}{300} = 0.17,$

E(XY)=$\sum_{x=1}^{3} \sum_{y=0}^{2} xyp(x,y) = 2 \cdot \frac{9}{300} + 3 \cdot \frac{36}{300} + 6 \cdot \frac{3}{300} = \frac{144}{300}$=0.48,

Cov(X,Y)=E(XY)-E(X)E(Y)=0.48-(2.64)(0.17)=0.0312>0.

Hence, *X and Y are positively correlated*.

~~Example 10.11 on P.442~~ (Skip) 題意不清楚

(10.11) Ann cuts an ordinary deck of 52 cards and displays the exposed card. Andy cuts the remaining stack of cards and displays his exposed card. Counting jack, queen, and king as 11, 12, 13, let X and Y be the numbers on the cards that Ann and Andy expose, respectively. Find Cov(X,Y) and interpret the result.

~~Ans:~~ $p(x,y) = P(X = x, Y = y) = P(Y = y|X = x)P(X = x).$

$p_X(x) = \frac{1}{13}, 1 \leq x \leq 13. p_Y(y) = \frac{1}{13}, 1 \leq y \leq 13.$

$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{287}{6} - 7 \times 7 = -\frac{7}{6}.$

Example 10.12 on P.443-P.444

(10.12) Let X be the lifetime of an electronic system and Y be the lifetime of one of its components. Suppose that the electronic system fails if the component does (but not necessarily vice versa). Furthermore, suppose that the joint probability density function of X and Y (in years) is given by

$$f(x,y) = \frac{1}{49}e^{-y/7} \ \ if\ 0 \le x \le y < \infty, and\ 0\ elsewhere.$$

(a) Determine the expected value of the remaining lifetime of the component when the system dies.

Ans: E(Y-X)=$\int_0^{\infty}\int_0^{y}(y-x)f(x,y)dxdy$ =7.

(b) The covariance of X and Y is Cov(X,Y)=E(XY)-E(X)E(Y)=147-7×14=49.

Examples 10.13-10.15 on P.445

(10.13) Let X be the number of 6's in n rolls of a fair die. Find Var(X).

Ans: Let $X_i = 1\ if\ on\ the\ i\text{th roll the die lands 6, and 0 otherwise.}$

Then, $X = X_1 + X_2 + \cdots + X_n$. Since $\{X_i's\}$ are independent, thus,

Var(X)=$nVar(X_1) = n(E(X_1^2) - [E(X_1)]^2) = \frac{5n}{36}$.

(10.14) Let X be the sum of n independent Bernoulli trials with the probability of success p, then $X \sim b(n,p)$, is a binomial random variable with parameters (n,p), thus, Var(X)=$np(1-p)$.

(10.15) Let X be the sum of n independent negative binomial r.v.s with parameters (r,p), then Var(X)=$n\frac{r(1-p)}{p^2}$.

Lemma 10.1 on P.446 and (A12) on P.447

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a distribution F with mean μ and variance σ^2. Let $\bar{X}$ be the sample mean. Then

$$E(\bar{X}) = \mu, \text{ and } Var(\bar{X}) = \frac{\sigma^2}{n}.$$

(A12) Let X and Y be coordinates of a random point selected uniformly from the unit disk $\{(x, y): x^2 + y^2 \leq 1\}$.

(a) Are X and Y independent? (No) $f(x, y) \neq f_X(x) f_Y(y)$

(b) Are they uncorrelated? (Yes) Cov(X,Y)=E(XY)-E(X)E(Y)=0

10.3 Correlation on P.450

(Definition 10.2) Let X and Y be two random variables with $0 < \sigma_X^2, \sigma_Y^2 < \infty$. The covariance between the standardized X and the standardized Y is called the *correlation coefficient* between X and Y,

denoted by $\rho(X,Y)$ which is given by ρ(X,Y)=$\frac{Cov(X,Y)}{\sigma_X \sigma_Y}$.

(Lemma 10.2) For random variables X and Y with correlation coefficient ρ(X,Y),

$$Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 2 + 2\rho(X,Y)$$

$$Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2 - 2\rho(X,Y)$$

Theorem 10.5 on P.452 and Proof of (a)

(Theorem 10.5) For random variables X and Y with correlation coefficient $\rho(X,Y)$, we have

(a) $-1 \leq \rho(X,Y) \leq 1$.

(b) With probability 1, $\rho(X,Y)$=1 iff $Y = aX + b$ for some constants $a, b, where\ a > 0$.

(c) With probability 1, $\rho(X,Y)$=-1 iff $Y = aX + b$ for some constants $a, b, where\ a < 0$.

(Proof of (a)) $Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$, by Theorem 10.3

$|Cov(X,Y)| \leq \sqrt{E[(X-E(X))^2]E[(Y-E(Y))^2]} = \sqrt{Var(X)Var(Y)} = \sigma_X\sigma_Y,$

thus, $|\rho(X,Y)| = |\frac{Cov(X,Y)}{\sigma_X\sigma_Y}| \leq 1, which\ completes\ the\ proof.$

Theorem 10.5(b) and Proof

(10.5(b)) For random variables X and Y with correlation coefficient $\rho(X, Y)$,

(b) With probability 1, $\rho(X, Y)$=1 iff $Y = aX + b$ for some constants $a > 0$, b.

(Proof of (b)) If $\rho(X, Y)$=1, by (Lemma 10.2) as given before,

$Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2 - 2\rho(X, Y) = 0$, then $\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = \text{c}$, a constant. Hence,

we have $Y = \frac{\sigma_Y}{\sigma_X}\text{X} - \text{c}\sigma_Y = aX + b, where\ b = -\text{c}\sigma_Y$, and $a = \frac{\sigma_Y}{\sigma_X} > 0$.

On the other hand, if $Y = aX + b$ with $a > 0, we\ have$

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{Cov(X, aX + b)}{\sigma_X \sigma_{aX+b}} = \frac{aCov(X, X)}{\sigma_X (a\sigma_X)} = 1.$$

Theorem 10.5 on P.452 and Proof of (c)

(Theorem 10.5) For random variables X and Y with correlation coefficient $\rho(X,Y)$, we have

(a) $-1 \leq \rho(X,Y) \leq 1.$

(b) With probability 1, $\rho(X,Y)$=1 iff $Y = aX + b$ for some constants $a, b, where\ a > 0.$

(c) With probability 1, $\rho(X,Y)$=-1 iff $Y = aX + b$ for some constants $a, b, where\ a < 0.$

(Proof of (c)) Similar to the proof of (b), we can prove (c) by using

$$Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 2 + 2\rho(X,Y) = 0\ if\ \rho(X,Y) = -1.$$

Example 10.16 on P.452-P.453

(Example 10.16) Show that if X and Y are continuous random variables with the joint pdf $f(x,y) = x + y \;\; if \; 0 < x, y < 1, \;\; and \; 0 \; elsewhere,$ then X and Y are not linearly related.

(Hint): Show that $\rho(X,Y) \neq 1 \; or \; -1.$

$f_X(x) = x + \frac{1}{2}, f_Y(y) = y + \frac{1}{2}, \;\; 0 < x, y < 1, then$

$E(X) = \int_0^1 x(x + \frac{1}{2})dx = \frac{7}{12}, \;\; E(X^2) = \frac{5}{12}, \;\; and \; \sigma_X = \frac{\sqrt{11}}{12}.$ Similarly,

$E(Y) = \frac{7}{12}, E(Y^2) = \frac{5}{12}, \;\; and \; \sigma_Y = \frac{\sqrt{11}}{12}.$

$E(XY) = \int_0^1 \int_0^1 (xy)(x + y)dxdy = \frac{1}{3}.$ Then

$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{7}{12} \times \frac{7}{12} = -\frac{1}{144},$ and $\rho(X,Y) = -\frac{1}{11}.$

Example 10.17 on P.453

(10.17) Let X be a random number from the interval (0, 1) and $Y = X^2$. The probability density function of X is $f(x) = 1 \; if \; 0 < x < 1$,

0 elsewhere, and for $n \geq 1$,

$$E(X^n) = \int_0^1 x^n \, dx = \frac{1}{n+1}.$$

$$\sigma_X^2 = \mathrm{Var(X)} = \frac{1}{12}, \; \sigma_Y^2 = \mathrm{Var(Y)} = \frac{4}{45},$$

$$Cov(X, Y) = \frac{1}{12},$$

$$\rho(X, Y) = \frac{\sqrt{15}}{4} = 0.968.$$

(Challenge) What do you conclude if $X \sim N(0,1)$?