Stat 567
2-9-17

Theorem: If X and Y are independent, then $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ ①

Proof: (Continuous case)

$$E[g(X)h(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y) f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y) \overbrace{f_X(x) f_Y(y)}^{\text{by independence}}\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\left[h(y)f_Y(y)\int_{-\infty}^{\infty} g(x) f_X(x)\,dx\right] dy$$

②

$$= E[g(X)] \int_{-\infty}^{\infty} h(y) f_Y(y)\,dy$$

$$= E[g(X)]\,E[h(Y)].$$

Defn: The covariance of X & Y is

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

Note: $E[XY - \mu_X Y - X\mu_Y + \mu_X\mu_Y]$

$$= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X\mu_Y$$

(3)

$$= E[XY] - \mu_X \mu_Y = E[XY] - E[X]E[Y]$$

Note: If X and Y are independent,

then $E[XY] = E[X]E[Y]$

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

The <u>converse</u> <u>is</u> <u>not</u> <u>true</u>

(4)

Properties of the covariance:

(1) $Cov(X,X) = E[X^2] - (E[X])^2$
$= V[X]$

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

(3) $Cov(cX,Y) = E[cXY] - E[cX]E[Y]$
$= cE[XY] - cE[X]E[Y]$
$= c\,Cov(X,Y)$

(4) $\mathrm{Cov}(X, Y+Z) = E[X(Y+Z)] - E[X]E[Y+Z]$

$= E[XY + XZ] - E[X](E[Y] + E[Z])$

$= \underline{E[XY]} + \underline{\underline{E[XZ]}} - \underline{E[X]E[Y]} - \underline{\underline{E[X]E[Z]}}$

$= \mathrm{Cov}(X,Y) + \mathrm{Cov}(X,Z)$

Note: The covariance is a "bilinear" operator

Find $V[\sum_{i=1}^{n} X_i]$

$= \mathrm{Cov}[\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j]$ now use (4) and (2)

$= \sum_{i=1}^{n} \mathrm{Cov}(X_i, X_i) + \sum\sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$

$= \sum_{i=1}^{n} V[X_i] + 2\sum\sum_{i<j} \mathrm{Cov}(X_i, X_j)$

(7)

Note: If X_i, X_j are independent $\forall i \neq j$,

then $V\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} V[X_i]$

Consequence: <u>Recall</u> $X \sim Bino(n,p)$

$X = \sum_{i=1}^{n} X_i$ where $X_i \sim Bern(p)$ and the X_i's are independent

So $V[X] = \sum_{i=1}^{n} V[X_i] = npq$

(8)

<u>Hypergeometric Distribution</u>

You have a population of N items

R of them are of a particular type.

Randomly select n items.

$X = \#$ items of that type in the sample.

$$p(x) = \frac{\binom{R}{x}\binom{N-R}{n-x}}{\binom{N}{n}} \quad x = 0,1,\ldots,\min(R,n)$$

9

Let $X_i = \begin{cases} 1 & \text{If the } i^{th} \text{ item is of that type} \\ 0 & \text{otherwise} \end{cases}$

$P[X_i = 1] = \frac{R}{N}$ Call this p

True $\forall i$ because the X_i's are "interchangeable"

$\therefore E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = np$

$= n\frac{R}{N}.$

10

$X_i \sim Bern(p)$

$V[X_i] = pq$

$V[X] = V[\sum_{i=1}^{n} X_i] = \underbrace{\sum_{i=1}^{n} V[X_i]}_{npq} + 2\sum\sum_{i<j} Cov(X_i, X_j)$

$Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$

$= E[X_i X_j] - p^2$

(11)

$$E[X_iX_j] = 0\cdot P[X_iX_j=0] + 1\cdot P[X_iX_j=1]$$

$$= P[X_iX_j=1]$$

$$= P[\; X_i=1 \text{ and } X_j=1]$$

$$= P[X_i=1]\cdot P[X_j=1 \mid X_i=1]$$

$$= p\cdot \frac{R-1}{N-1}$$

So $\mathrm{Cov}(X_i,X_j) = p\,\frac{R-1}{N-1} - p^2$

(12)

And $V[X] = npq + 2\binom{n}{2}\left(p\,\frac{R-1}{N-1} - p^2\right)$

$$= npq + \cancel{2}\,\frac{n!}{\cancel{2}!\,(n-2)!}\,p\left(\frac{R-1}{N-1} - p\right)$$

$$= npq + n(n-1)p\underbrace{\left(\frac{R-1}{N-1} - \frac{R}{N}\right)}_{\frac{\cancel{NR}-N-\cancel{NR}+R}{N(N-1)}}$$

$$= npq + n(n-1)p\,\frac{\left(-1+\frac{R}{N}\right)}{N-1}$$

(13)

$$= npq - \frac{n(n-1)pq}{N-1}$$

$$= npq\left[\underbrace{1 - \frac{(n-1)}{N-1}}_{\frac{N-1-n+1}{N-1}}\right]$$

$$V[X] = npq\underbrace{\left(\frac{N-n}{N-1}\right)}_{\text{f.p.c.}} = n\frac{R}{N}\left(1-\frac{R}{N}\right)\left(\frac{N-n}{N-1}\right)$$

f.p.c. = finite population correction

(14)

Defn: Suppose $X_1, \ldots, X_n$ are $\underbrace{\text{i.i.d.}}_{\text{independent, identically distributed}}$

with mean μ and variance σ^2.

Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then $\bar{X}$ is called the sample mean.

Properties: ① $E[\bar{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$

$$= \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu$$

(15)

(2) $V[\bar{X}] = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} V[X_i] + 0\right]$$

$$= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

(3) $Cov(\bar{X}, X_i - \bar{X}) = Cov(\bar{X}, X_i) - \underbrace{Cov(\bar{X}, \bar{X})}_{\substack{V[\bar{X}] \\ \sigma^2/n}}$

$$= Cov\left(\frac{1}{n}\sum_{i=1}^{n} X_i, X_i\right) - \frac{\sigma^2}{n}$$

(16)

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

$$= \frac{1}{n}\underbrace{Cov(X_i, X_i)}_{\substack{V[X_i] \\ \sigma^2}} - \frac{\sigma^2}{n} = 0$$

So $Cov(\bar{X}, X_i - \bar{X}) = 0 \quad \forall i$

Tuesday: Midterm exam

1-page of notes (front and back)

Calculator

Covers everything from the beginning thru Thurs Feb 7

53. If X is uniform over $(0, 1)$, calculate $E[X^n]$ and $\mathrm{Var}(X^n)$.

55. Suppose that the joint probability mass function of X and Y is

$$P(X = i, Y = j) = \binom{j}{i} e^{-2\lambda}\lambda^j / j!, \quad 0 \le i \le j$$

(a) Find the probability mass function of Y.
(b) Find the probability mass function of X.
(c) Find the probability mass function of $Y - X$.

61. Let X and W be the working and subsequent repair times of a certain machine. Let $Y = X + W$ and suppose that the joint probability density of X and Y is

$$f_{X,Y}(x, y) = \lambda^2 e^{-\lambda y}, \quad 0 < x < y < \infty$$

(a) Find the density of X.
(b) Find the density of Y.
(c) Find the joint density of X and W.
(d) Find the density of W.

76. Let X and Y be independent random variables with means μ_x and μ_y and variances σ_x^2 and σ_y^2. Show that

$$\mathrm{Var}(XY) = \sigma_x^2\sigma_y^2 + \mu_y^2\sigma_x^2 + \mu_x^2\sigma_y^2$$

79. With $K(t) = \log(E\left[e^{tX}\right])$, show that

$$K'(0) = E[X], \quad K''(0) = \mathrm{Var}(X)$$