20249433 최재필

$E(X)=\mu$, $E(X^2)=\sigma^2+\mu^2$
$E(\bar{X})=\mu$, $E(\bar{X}^2)=\frac{\sigma^2}{n}+\mu^2$
$E(X_i-\bar{X})=E(X_i)-E(\bar{X})=0$

HW 3

1. If $X_1, X_2, \dots, X_n$ are i.i.d. random variables with mean μ and variance σ^2, calculate the covariance $\bar{X}$ and $X_i - \bar{X}$ for any $i = 1, \dots, n$.

$$Cov(\bar{X}, X_i-\bar{X}) = E[(\bar{X}-\underbrace{E(\bar{X})}_{=\mu})((X_i-\bar{X})-\underbrace{E(X_i-\bar{X})}_{=0})] = E[(\bar{X}-\mu)(X_i-\bar{X})] = E[\bar{X}\cdot X_i - \bar{X}^2 - \mu X_i + \mu\bar{X}]$$

$$= E[\bar{X}\cdot X] - E[\bar{X}^2] - \cancel{\mu \underbrace{E[X_i]}_{=\mu}} + \cancel{\mu \underbrace{E[\bar{X}]}_{=\mu}} = E[\frac{1}{n}\sum_{j=1}^{n} X_j \cdot X_i] - (\frac{\sigma^2}{n}+\mu^2) = \frac{1}{n}\cdot\sum_{j=1}^{n}E(X_j\cdot X_i) - \frac{\sigma^2}{n}+\mu^2$$

when $i\neq j$, X_i and X_j are independent
$\because$ iid $E(X_i\cdot X_j)=E(X_i)\cdot E(X_j)$
when $i=j$, $E(X_i\cdot X_j)=E(X_i^2)$

$$= \frac{1}{n}\left(\underset{i=j}{E(X_i^2)} + \sum_{\substack{j=1 \\ i\neq j}}^{n} E(X_i)\cdot E(X_j)\right) - \frac{\sigma^2}{n}+\mu^2 = \frac{1}{n}(\sigma^2+\mu^2 + (n-1)\cdot\mu^2) - \frac{\sigma^2}{n}+\mu^2$$

$$= \frac{1}{n}(\sigma^2+n\cdot\mu^2) - \frac{\sigma^2}{n} - \mu^2 = 0$$

2. A commercial for a manufacturer of household appliance claims that 3% of all its product require a service call in the first year. A consumer protection association wants to check the claim by surveying 400 households that recently purchased one of the company's appliances. What is the probability that more than 5% require a service call within the first year?

(a) Calculate the probability using the approximate normal approach.

$n=400$
$p=3\%$

$X\sim B(400, 3\%)$
approx $\sim N(12, 11.64)$

$P(X>20) = 0.009517$

when $p=5\%$, about 20 households require a service.

(b) Calculate the probability using the binomial distribution.

$X\sim B(400, 3\%)$ $\quad$ $P(X>20) = 0.01045$

3. Suppose that X has normal distribution with mean μ = 10 and standard deviation σ = 2. Doing the following parts:

$X\sim(10, 2^2)$

(a) Calculate $P(6 < X < 14)$.

$\|$

$P(-2<Z<2) = 0.9545$

(b) Find the percentile c such that P(X ≤ c) = 0.95.

$$\frac{C-10}{2} = 1.64487, \quad C = 13.2897$$

$n=4$ $\qquad X_i \sim N(10, 2^2)$

(c) A random sample of size 4, {X1, ...,X4}, is taken from the normal distribution with mean = 10 and standard deviation = 2. Find the probability that the average of this sample is at most 12.

$= \bar{X}.$

$$P(\bar{X} \le 12) = 0.97724$$

$$\bar{X} \sim N(10, \underbrace{\frac{2^2}{4}}_{=1}) \quad \because \text{if } X_i \sim N, \ \sum_{i=1}^{4} X_i \sim N$$

(NOT by CLT)

4. Bits are sent over a communications channel in packets of 160. If the probability of a bit being corrupted (one error) over this channel is 0.2 and such errors are independent. Let X denotes the number of bits that are corrupted over this channels.

$p=$ $\quad n=$

(a) What is the distribution of X? Can it be approximated as normal distribution?

$$E(X) = np = 32$$
$$Var(X) = npq = 25.6$$

Yes. $\quad X \sim Binom(160, 0.2)$

$$\underset{approx.}{\sim} N(32, 25.6)$$

(b) Approximately, what is the probability that more than 50 bits in a packet are corrupted?

$$P(X \ge 50) = P(Z \ge 3.5575) = 0.0001871$$

$\overset{=}{\to}$ sampling with replacement

5. A large population is described by the probability distribution

x	f(x)
0	0.1
1	0.2
2	0.7

Let X_1 and X_2 be a random sample of size 2 from the distribution

(a) Determine the sampling distribution of $\max(X_1,X_2)$?

max(X1,X2)	P(X1, X2)
0	0.01
1	0.08
2	0.91

(b) Determine the sampling distribution of X_1+X_2?

X1+X2	P(X1, X2)
0	0.01
1	0.04
2	0.18
3	0.28
4	0.49

$E(X_i)=\mu,\ Var(X_i)=\sigma^2$

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

$E(X_i-\bar{X})=E(X_i)-E(\bar{X})=\mu-\mu=0$

$Var(X_i-\bar{X})=Var(X_i)-2Cov(X_i,\bar{X})+Var(\bar{X})$

$=$

$Cov(X_i,\bar{X})=E(X_i\cdot\bar{X})-E(X_i)\cdot E(\bar{X})=E(X_i\cdot\bar{X})-\mu^2$

$E(X_i^2)=\sigma^2+\mu^2$

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

$* \ E[X_i\cdot X_j]=$

- $i\neq j$ i.i.d이므로 X_i와 X_j가 독립. $=E(X_i)\cdot E(X_j)=\mu\cdot\mu=\mu^2$
- $i=j$ $X_i=X_j$ $=E[X_i^2]=\sigma^2+E(X_i)^2=\sigma^2+\mu^2$

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

$=\frac{1}{n}E\left[\sum_{j=1}^{n}X_iX_j\right]$

$=\frac{1}{n}\cdot\sum_{j=1}^{n}E[X_iX_j]=\frac{1}{n}\sum_{\substack{j=1\\ i\neq j}}^{n}E[X_i]\cdot E[X_j]+\frac{1}{n}\underset{i=j}{E[X_j^2]}=\frac{n-1}{n}\cdot\mu^2+\frac{\sigma^2+\mu^2}{n}$

$=\frac{n\cdot\mu^2-\mu^2+\sigma^2+\mu^2}{n}=\frac{\sigma^2}{n}+\mu^2$

?

$* \ E\left[\sum_{i=1}^{n}X_i\right]=\sum_{i=1}^{n}E[X_i]$

proof

$\sum_{i=1}^{n}E[X_i]=\sum_{i=1}^{n}\cdot\sum_{j=1}^{m}x_j\cdot P(X_i=x_j)=$

(loop)(loop)

m

$x_1\cdot P(X_1=x_1)+x_2\cdot P(X_1=x_2)+\dots+x_m\cdot P(X_1=x_m)$

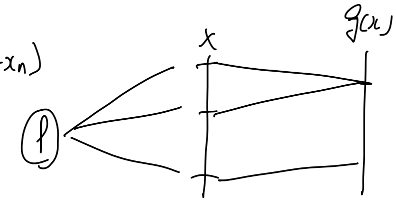
n

$x_1\cdot P(X_2=x_1)+\cdots\quad +x_m\cdot P(X_n=x_m)$

$\vdots$

$E\left[\sum_{i=1}^{n}X_i\right]=E[X_1+X_2+X_3+\cdots+X_n]$

$=\sum_{j=1}^{m}P(X_1=x_j,X_2=x_j,\dots,X_n=x_j)\cdot(x_1+x_2+\cdots+x_n)$



linearity of expectation

$$\mathrm{Cov}(X, \bar{X}) = E\big((X - E[X]) \cdot (\bar{X} - E[\bar{X}])\big) = E[(X-\mu)(\bar{X}-\mu)]$$

$E[X_i] = \mu$

$E[\bar{X}] = \mu$

$$= E\left[(X-\mu)\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right)\right] = E\left[X \cdot \frac{1}{n}\sum_{i=1}^{n} X_i - \mu X - \frac{\mu}{n}\sum_{i=1}^{n} X_i + \mu^2\right]$$

$$= E\left[X \cdot \frac{1}{n}\sum_{i=1}^{n} X_i\right] - \cancel{\mu^2} \quad \cancel{+\mu^2} \quad - \frac{\mu}{n} E\left[\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{n} E\Big[\underbrace{X_i^2}_{i=j} + \underbrace{X_i(X_1 + X_2 + \cdots + X_n)}_{i \neq j}\Big] - \frac{\mu}{n} \cdot n\mu$$

$$= \frac{1}{n}\Big(E(X^2) + \sum_{i \neq j} E(X_i X_j)\Big) - \mu^2 = \frac{1}{n}\big(\sigma^2 + \mu^2 + (n-1)\mu^2\big) - \mu^2$$

X_i, X_j iid independent

$$= \frac{1}{n}(\sigma^2 + n\mu^2) - \mu^2$$

$$= \frac{\sigma^2}{n} + \mu^2 - \mu^2 = \frac{\sigma^2}{n}$$

$$E(X_i \cdot \bar{X}) = E\left(X_i \cdot \sum_{j=1}^{n} \frac{X_j}{n}\right) = \frac{1}{n} \cdot E\left(X_i \cdot \sum_{j=1}^{n} X_j\right) = \frac{1}{n} E\left(\underset{i=j}{X_i^2} + \sum_{\substack{j=1 \\ i \neq j}}^{n} X_i \cdot X_j\right) = \frac{1}{n} \cdot E(X_i^2) + \frac{1}{n} \cdot (n-1) \cdot E(X_i) \cdot E(X_j)$$

$$= \frac{1}{n} \cdot (\sigma^2 + \mu^2) + \frac{1}{n} (n-1) \cdot \mu^2$$

$$= \frac{\sigma^2 + \cancel{\mu^2} + n\mu^2 - \cancel{\mu^2}}{n} = \frac{\sigma^2}{n} + \mu^2$$

틀린건 아니지만, 이걸 필요 없음 $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$가 되니까

$\bar{X}$ $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

NOT by CLT.

갈께.

(c) A random sample of size 4, {X1, ...,X4}, is taken from the normal distribution with mean = 10 and standard deviation = 2. Find the probability that the average of this sample is at most 12.

$X \sim N(10, 2^2)$ $\bar{X}$ $n=4$

$X \sim N$ 이니까

$\sum_{i=1}^{4} X_i$ 도 $\sim N$ 이기 때문

$$P\left(\frac{X_1+X_2+X_3+X_4}{4} \leq 12\right) = P(X_1+X_2+X_3+X_4 \leq 48) = 0.97724$$

$E(X_1+X_2+X_3+X_4) = 10+10+10+10 = 40$ $(X_1+X_2+X_3+X_4) \sim N(40, 16)$

$Var(X_1+X_2+X_3+X_4) = 4+4+4+4 = 16$

$\because X_1, X_2, X_3, X_4$ are all independent