중간고사 대체 과제

1. $G = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$

$P(\omega_1) = P(\omega_4) = P(\omega_5) = \alpha$

$P(\omega_2) = 2P(\omega_1) = 2\alpha$

$P(\omega_3) = 4P(\omega_1) = 4\alpha$

$\because P(G) = 1 = 9\alpha \quad \therefore \alpha = 1/9 \quad \therefore P(\omega_1) = P(\omega_4) = P(\omega_5) = \frac{1}{9}$

$P(\omega_2) = \frac{2}{9}, \; P(\omega_3) = \frac{4}{9}$

$A = \{\omega_1, \omega_4\}$

$P(A) = P(\omega_1) + P(\omega_4) = \frac{2}{9} \qquad \therefore P(A) = \frac{2}{9}$

2. 남자 집단. 여자집단에서의 왼손잡이를 X, Y라고 할때, $p_1 = 0.043$, $p_2 = 0.036$ 이고

표본비율 $\hat{p}_1$, $\hat{p}_2$ 는 중심극한 정리에 의해 $N(p_1, \frac{p_1(1-p_1)}{n_1})$, $N(p_2, \frac{p_2(1-p_2)}{n_2})$ 를 따른다.

$\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}) = N(0.007, 0.000169062)$

비율의차이 $P(\hat{p}_1 - \hat{p}_2 \le 0.005) = P(\frac{\hat{p}_1 - \hat{p}_2 - 0.007}{0.000169062} \le \frac{0.005 - 0.007}{0.000169062}) = P(Z \le -0.1538) = 0.4404$

$\therefore 0.4404$

3. 1) 18, 19, 20, 21, 21, 22, 22, 23, 23

24, 24, 24, 25, 25, 26, 26, 27, 28

29, 30

2) $Q_1 = \quad p = 0.25, \; n = 20 \rightarrow np = 5 \qquad x_{(5)} = 21$

$Q_2 = \quad p = 0.5, \; n = 20 \rightarrow np = 10 \qquad \frac{x_{(10)} + x_{(11)}}{2} = \frac{24+24}{2} = 24$

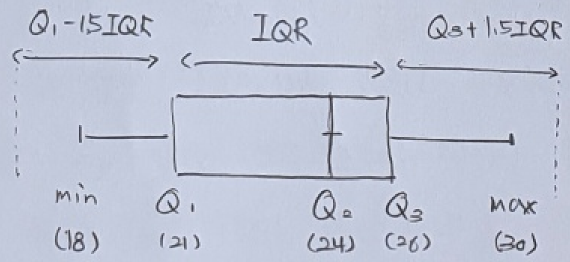
$Q_3 = \quad p = 0.75 \quad n = 20 \rightarrow np = 15 \qquad x_{(15)} = 26$

$IQR = Q_3 - Q_1 = 26 - 21 = 5$

좌측경계 : $Q_1 - IQR \times 1.5 = 21 - 7.5 = 13.5$

우측경계 : $Q_3 + 1.5IQR = 26 + 7.5 = 33.5$

$\therefore$



max : 30

Q_3 : 26

Q_2 : 24

Q_1 : 21

min : 18

3) ∴ 줄기-잎 도표의 경우에는 자료구조의 형태와 자료의 정보를 상실하지 않고

원자료값들 알 수있어 정보상실의 부담을 덜 수 있다.

위데이터 의 경우, 중앙값을 중심으로 (대체적으로) 좌우 대칭의 형태를 띄고 있어, 정규분포의

형태를 가지고 있다고 해석가능하다. 이에 따라 중앙값 ≒ 평균의 결과도 생각가능하다

4. $n=9$ 　 재점탈 : 34 62 65 67 69 69 70 72 99

$\alpha = 0.15$

$[n\alpha] = 1$

$$\bar{x}_{0.15} = \frac{x_{([n\alpha]+1)} + \cdots + x_{(n-[n\alpha])}}{n-2[n\alpha]}$$

$$= \frac{x_2 + x_3 + x_4 + \cdots + x_8}{9-2} = \frac{62+65+67+69+69+70+72}{7}$$

$$= \frac{474}{7} = 67.7143$$

$\therefore \bar{x}_{0.15} = 67.7143$

5 1) $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ 　 $n=3,\ p=0.6$

$\therefore$

x	$f(x)$	$xf(x)$	$x^2f(x)$
0	$\binom{3}{0}(0.6)^0(0.4)^3 = 0.064$	0	0
1	$\binom{3}{1}(0.6)(0.4)^2 = 0.288$	0.288.	0.288
2	$\binom{3}{2}(0.6)^2(0.4) = 0.432$	0.864	1.728
3	$\binom{3}{3}(0.6)^3 = 0.0216$	0.648	1.944
합계	1	1.8	3.96

2) $E(x) = \Sigma x f(x) = 0.288 + 0.864 + 0.648 = 1.8$

$Var(x) = \Sigma x^2 f(x) - (E(x))^2 = 3.96 - (1.8)^2 = 0.72$, $sd(x) = \sqrt{Var(x)}$

$\therefore E(x) = 1.8,\ sd(x) = 0.8485$

3) $E(x) = np = 1.8$

$Var(x) = npq = 1.8 \times 0.4 = 0.72.$

$sd(x) = \sqrt{Var(x)} = 0.8485$ 　 $\therefore$ (2)의 결과와 동일하다.

6. $\therefore$ 중심극한정리는 분포에 상관없이 n이 크면 표본평균 $\bar{X}$의 분포는 근사적으로 정규분포가 $N(\mu, \frac{\sigma}{\sqrt{n}})$ 된다. 따라서 $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ 는 근사적으로 표준정규분포 $N(0,1)$을 따른다. 이에따라

n이 큰 어떠한 모든 분포에 대해서 확률을 이용한 수학적 판단이 가능해진다는 이점이 있을 것이다.

7. $f(x) = \binom{30}{x}(0.2)^x(0.8)^{30-x}$

$P(X=4) = \binom{30}{4}(0.2)^4(0.8)^{26} = 0.1325$

$\mu = np = 6,\ \sigma^2 = \sqrt{npq}^2 = 4.8 \rightarrow X \simeq (6.4.8)$ 이때 $\overset{(4-0.5)}{3.5}$와 $\overset{(4+0.5)}{4.5}$를 표준화하면.

$P(X=4) : P(3.5 \le X \le 4.5) = P\left(\frac{3.5-6}{\sqrt{4.8}} \le x \le \frac{4.5-6}{\sqrt{4.8}}\right) \simeq P(-1.1410 \le Z \le -0.68)$

$= P(Z \le -0.68) - P(Z \le -1.141) = 0.2483 - 0.1271 = 0.1212$ $\therefore$ pmf $P(X=4) = 0.1325$

수정하면 $P(X=4) = 0.1212$

8. $P(M) = \frac{40}{60} = \frac{2}{3}$, $P(A) = \frac{38}{60} = \frac{19}{30}$

$P(M, A) = \frac{24}{60} = \frac{2}{5}$

$\because P(MA) \ne P(M)P(A)$ $P(MA) = \frac{2}{5}$, $P(M)P(A) = \frac{19}{45}$

$\therefore$ case M.A are not independent

9. $d = 2.5,\ \alpha = 1 - 0.9 = 0.1$

$n \ge \left[\frac{Z_{\alpha/2}\sigma}{d}\right]^2$ $Z_{\alpha/2} = Z_{0.05} \longrightarrow P[Z \le Z_{0.05}] = 0.05$

$\therefore Z_{0.05} = 1.645$

$110 = \left[\frac{1.645 \cdot \sigma}{2.5}\right]^2$ $\sigma = \frac{d\sqrt{n}}{Z_{0.05}} = \frac{2.5\sqrt{110}}{1.645} = 15.9393$

if $\alpha = 1 - 0.95 = 0.05$, $d = 2$

$n \ge \left[\frac{Z_{0.025} \cdot \sigma}{d}\right]^2 = \left[\frac{1.96 \times 15.9393}{2}\right]^2 = 244.0004$

$\therefore n = 245$

10. $\therefore$ 신뢰구간은 표본에서 추출되어 알 수없는 모집단 모수 값이 포함될 가능성이 있는 범위이다. 이때 신뢰구간 95%는 표본중 95%가 모집단 모수를 포함하는 구간을 나타냄을 의미한다. 예를들어 $\bar{X}$의 평균이 μ, 표준편차가 $\frac{\sigma}{\sqrt{n}}$ 일때, 표준정규 분포에 의해 확률변수가 $1.96\frac{\sigma}{\sqrt{n}}$에 있을 확률이 0.95 라면.

$P[\mu - 1.96\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1.96\frac{\sigma}{\sqrt{n}}]$와 같이 나타낼수 있다.

이때 μ와 $\bar{X}$를 정리하면

$P[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}]$ 이고 확률구간이 모수 μ를 포함할 확률이 0.95가 된다.

11. A : 실제로 불량일 사상 , D : 불량으로 판정될 사상

$P(A) = 0.06$. $P(A^C) = 0.94$.

$P(D|A^c) = 0.03$, $P(D^c|A) = 0.02$, $P(D|A) = 1 - P(D^c|A) = 0.98$

$D = (A \cap D) \cup (A^c \cap D)$ $A \cap D$, $A^c \cap D$: 상호배반

$$P(D) = P(A \cap D) \cup P(A^c \cap D) = P(A)P(D|A) + P(A^c)P(D|A^c)$$
$$= (0.06 \times 0.98) + (0.94 \times 0.03)$$
$$= 0.0588 + 0.0282 = 0.087$$

$\therefore 0.087$

12 1) ∴

x	$f(x)$	$xf(x)$	$x^2f(x)$
0	$\frac{1}{84}\binom{5}{0}\binom{4}{3} = \frac{4}{84}$	0	0
1	$\frac{1}{84}\binom{5}{1}\binom{4}{2} = \frac{30}{84}$	$\frac{30}{84}$	$\frac{30}{84}$
2	$\frac{1}{84}\binom{5}{2}\binom{4}{1} = \frac{40}{84}$	$\frac{80}{84}$	$\frac{160}{84}$
3	$\frac{1}{84}\binom{5}{3}\binom{4}{0} = \frac{10}{84}$	$\frac{30}{84}$	$\frac{90}{84}$
합계	1	$\frac{140}{84}$	$\frac{280}{84}$

2) $E(X) = \sum xf(x) = \frac{140}{84} = 1.6667$

$Var(x) = \sum x^2 f(x) - E(x)^2 = \frac{280}{84} - (1.6667)^2 = 0.5554$

$sd(x) = \sqrt{var(x)} = 0.7453$

$\therefore E(X) = 1.667$, $sd(x) = 0.7453$

13. ∴ P-값은 주어진 관측값을 이용하여 H_0를 기각할 수 있는 최소의 유의수준으로.
P-값이 유의수준 α보다 작다면 α가 P-값 보다 작게될때까지 H_0를 기각하고
H_1을 지지한다, 반대의 경우는 (P-val > 유의수준 α) 유의수준 α에 대하여
H_0을 기각하지 못하고 H_1을 지지하지 못하게 된다.

14. 성공확률 : $P = 0.7$

12번 이상 성공 ⇒ 11번까지 9번 이하로 성공한 상태 $P[X \le 9]$

확률변수 : $n = 11$, $P = 0.7$ 이항분포

$$f(x) = \binom{11}{x}(0.7)^x(0.3)^{11-x}$$

$P[X \le 9] = 0.887$ ∵ 누적 이항분포표 이용.

∴ 0.887.

15. $COV(X,Y) = E(XY) - E(X)E(Y)$

$= 10 - 3 \times 4 = -2$

$$\rho\,(\text{coefficient}) = \frac{COV(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{-2}{\sqrt{E(X^2)-E(X)^2}\sqrt{E(Y^2)-E(Y)^2}} = \frac{-2}{\sqrt{16-9}\sqrt{25-16}} = \frac{-2}{3\sqrt{7}}$$

∴ $COV(X,Y) = -2$, $\rho = -0.2520$. X, Y는 약한 음의 상관관계를 지닌다.

16. $X \sim N(341.7, 31.2)$, $P(\overline{X} < 313.4)$

$$P(\overline{X} < 313.4) = P\left(\frac{\overline{X} - 341.7}{31.2} < \frac{313.4 - 341.7}{31.2}\right)$$

$= P(Z < -0.9071) \fallingdotseq 0.1660$

∴ 0.1660

17. ① 귀무가설 : $H_0 = 100$.

대립가설 : $H_1 \ne 100$

② 검정통계량 : because $n > 30$, $\overline{X} \sim N(100, 5^2)$ $Z = \dfrac{100.78 - 100}{\frac{5}{10}} = 1.56$

③ 기각역 설정 : $H_1 \ne 100 \longrightarrow R: |Z| \ge Z_{0.025}$: $Z < -1.96$, $Z > 1.96$

∵ 검정통계량 1.56이 기각역을 만족하지 못하므로

귀무가설은 기각되지 않음, 대립가설을 지지하지 못한다.

∴ 평균은 변하지 않았다.

18. 음주운전 사고 확률 : 0.55

음주운전 아닌 사고 : 0.45

확률분포 X는 $n=5, p=0.55$ 인 이항분포를 따른다.

$$P(x) = \binom{5}{x}(0.55)^x(0.45)^{5-x}$$

(1) 5번 모두 음주운전 사고 $P(X=5) = \binom{5}{5}(0.55)^5(0.45)^0$

$= (0.55)^5 = 0.0313$ $\therefore 0.0313$

(2) 꼭 3번 사고 $P(X=3) = \binom{5}{3}(0.55)^3(0.45)^2 = 0.3369$ $\therefore 0.3369$

(3) 적어도 한번 $P(X \geq 1) = 1 - P(X < 0)$

$= 1 - \binom{5}{0}(0.55)^0(0.45)^5 = 0.9815$ $\therefore 0.9815$