#2.2

(1) 우선 수명의 줄기-잎 그림을 그리면

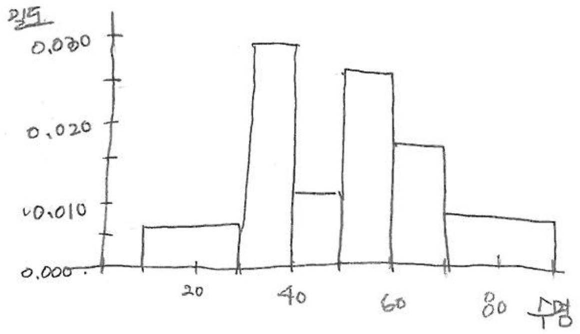
1 | 7
2 | 03
3 | 112234789
4 | 159
5 | 2345677
6 | 0378
7 | 4
8 | 3

이다.

이를 토대로 히스토그램을 위한 밀도와 계급구간을 나누면

구간	밀도	상대도수
10이상 30미만	0.00555	0.111
30이상 40미만	0.0297	0.297
40이상 50미만	0.0111	0.111
50이상 60미만	0.0259	0.259
60이상 70미만	0.0148	0.148
70이상 90미만	0.0037	0.074

이를 토대로 히스토그램을 그리면



이다. 수명은 두 그래프를 통해 30-60대 연령의 수명이 가장 많고, 30대, 50대의 수명을 가진 임금은 많지만 그 중간지점인 40대의 수명을 가진 임금은 상대적으로 적음을 알 수 있다.

그리고 재위기간의 줄기-잎 그림을 그리면

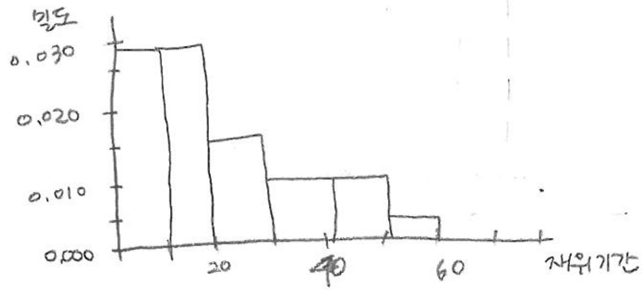
0 | 0.75 1.083
0 | 223347
1 | 024455558
2 | 24557
3 | 249
4 | 146
5 | 2

이다.

이를 토대로 밀도와 계급구간을 나누면

구간	밀도	상대도수
0이상 10미만	0.0296	0.296
10이상 20미만	0.0296	0.296
20이상 30미만	0.0148	0.148
30이상 40미만	0.0111	0.111
40이상 50미만	0.0111	0.111
50이상 60미만	0.0038	0.038

히스토그램을 그리면



이다. 두 그래프를 통해 재위기간은 0~20년이 가장 많고 재위기간이 긴 임금은 점점 그 수가 감소함을 알 수 있다.

(2) 우선 수명의 평균, 중앙값, 표준편차를 구하면

평균 : $(17+20+23+\cdots+83)\div 27 = 1269 \div 27 = 47$

중앙값 : $n=27$이므로 14번째 값

$\therefore 49$

표준편차 : $\sqrt{\frac{1}{26}\sum_{i=1}^{27}(x_i-47)^2}$

$\sum_{i=1}^{27} x_i = 1269 \quad \sum_{i=1}^{27} x_i^2 = 66985$ 이므로

$\sqrt{\frac{1}{26}\left(66985-\frac{1269^2}{27}\right)} = \sqrt{\frac{7342}{26}}$

$\fallingdotseq 16.804$

재위기간의 평균, 중앙값, 표준편차를 구하면

평균 : $(0.075+1.083+2+\cdots+52)\div 27 = 521.833 \div 27 = 19.327$

중앙값 : 14번째 값 : 15

표준편차 : $\sum_{i=1}^{27} x_i = 521.833 \quad \sum_{i=1}^{27} x_i^2 = 16279.179$

$\sqrt{\frac{1}{26}\sum_{i=1}^{27}(x_i-19.327)^2}$

$= \sqrt{\frac{1}{26}\left(16279.179-\frac{(521.833)^2}{27}\right)} \fallingdotseq 15.434$

(3) 상자그림을 그리기 위해 Q_1, Q_2, Q_3를 구하려면

n=27이므로 $26\times0.25+1 = 7.5$ (Q_1)

$26\times0.75+1 = 20.5$ (Q_3)

$Q_1 = \frac{33+34}{2} = 33.5$

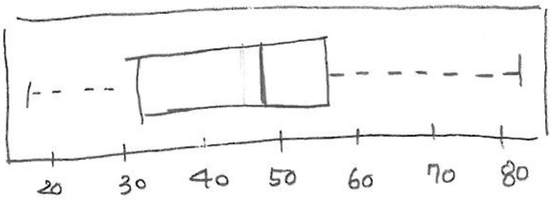
$Q_2 = 49$

$Q_3 = \frac{57+57}{2} = 57$

$IQR = 57-33.5 = 23.5$

$L = 33.5 - (1.5\times23.5) = -1.75$

$U = 57 + (1.5\times23.5) = 92.25$ 이므로 상자그림은



으로 나타낸다. 그러므로 이상점인 임금은 존재하지 않는다.

(4) 단종의 자료를 제외하면

수명 평균 : $(20+\cdots+83)\div26 = 1252\div26 = 48.154$

중앙값 : 13, 14번째 값의 평균 : $\frac{49+52}{2} = 50.5$

표준편차 : $\sum_{i=1}^{26} x_i = 1252$, $\sum_{i=1}^{26} x_i^2 = 66696$

$\sqrt{\frac{1}{25}\left(66696 - \frac{1252^2}{26}\right)} = \sqrt{\frac{6407.385}{25}} = 16.009$.

재위기간 평균 : $(0.075+1.083+\cdots+52)\div26 = 518.833\div26$

$= 19.955$

중앙값 : 13, 14번째 값의 평균 : $\frac{15+15}{2} = 15$

표준편차 : $\sum_{i=1}^{26} x_i = 518.833$, $\sum_{i=1}^{26} x_i^2 = 16270.179$

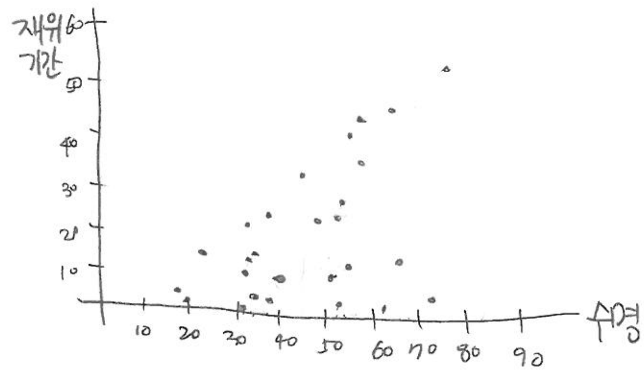
$\sqrt{\frac{1}{25}\left(16270.179 - \frac{518.833^2}{26}\right)} = \sqrt{\frac{5916.807}{25}}$

$= 15.384$

(2)와 비교했을때, 수명은 전체적으로 올라갔으나
재위기간은 비슷하고,
재위기간에는 영향을 미친다고 볼수 있다.

(5) 우선 수명을 x, 재위기간을 y축으로 해서

산점도를 그리면



이다. 여기서 피어슨 상관계수는

$$r_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{S_x}\right)\left(\frac{y_i-\bar{y}}{S_y}\right)$$

$$= \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}$$

$\bar{x} = 47$, $\bar{y} = 19.327$. $n=27$

$S_{xy} = \sum_{i=1}^{27} x_i y_i - 27\times47\times19.327$

$= (518+126+\cdots159) - 24525.963$

$= 28165.91 - 24525.963 =$

$= 3639.947$

$S_{xx} = 7342$

$S_{yy} = 6193.790$

$$r = \frac{3639.947}{\sqrt{7342}\sqrt{6193.790}} = \frac{3639.947}{6743.501} = 0.540$$

이므로 수명과 재위기간은

양의 상관관계를 가진다고 볼 수 있다.

(6) 양세계보의 내시 수명이 기록되어있는자는 약81명이다. 따라서 777명의 내시중 81명이 어느정도 표본으로서 적합한지 증명한다면, 임금의 수명과 내시의 수명의 비교가 가능하다. 또한 통계적으로 임금의 수명이 47세가 평균인데, 양세계보의 내시수명인 70과 비교하여 큰 차이가 나므로 설명 가능할것으로 보인다.

#2.7

(1) 첫째 머리길이 : x

첫째 머리폭 : y

둘째 머리길이 : z

둘째 머리폭 : w 라하면

$\sum x_i = 191+195+\cdots+190 = 4643$

$\sum y_i = 155+149+\cdots+163 = 3778$

$\sum z_i = 179+201+\cdots+187 = 4596$

$\sum w_i = 145+152+\cdots+150 = 3731$

$\sum x_i^2 = 191^2+195^2+\cdots+190^2 = 864585$

$\sum y_i^2 = 155^2+149^2+\cdots+163^2 = 572236$

$\sum z_i^2 = 847348$

$\sum w_i^2 = 557895$.

$\bar{x} = 4643/25 = 185.72$

$S_x = \sqrt{\frac{1}{24}\left(864585 - \frac{4643^2}{25}\right)} = 9.762$

$\bar{y} = 3778/25 = 151.12$

$S_y = \sqrt{\frac{1}{24}\left(572236 - \frac{3778^2}{25}\right)} = \sqrt{\frac{1304.64}{24}} = 7.373$

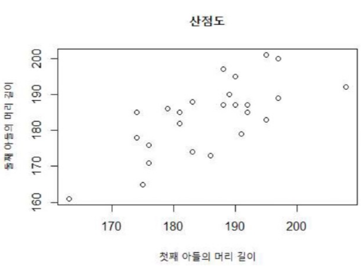
$\bar{z} = 4596/25 = 183.84$

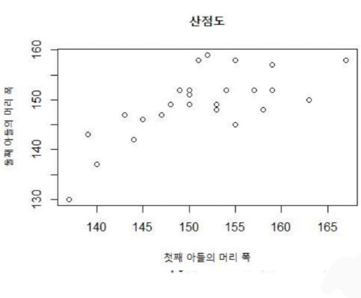
$S_z = \sqrt{\frac{1}{24}\left(847348 - \frac{4596^2}{25}\right)} = \sqrt{\frac{2419.36}{24}} = 10.040$

$\bar{w} = 3731/25 = 149.24$

$S_w = \sqrt{\frac{1}{24}\left(557895 - \frac{3731^2}{25}\right)} = \sqrt{\frac{1080.56}{24}} = 6.710$.

(2) ① 첫째 아들과 둘째 아들의 머리 길이와 폭에 대한 산점도를 그려라.





② 첫째 아들과 둘째 아들의 머리 길이와 폭에 대한 표본상관계수를 구하여라.

피어슨의 표본상관계수는 다음과 같은 식으로 표시된다.

$$R_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}$$

· 머리 길이

$$S_{xx} = \sum_{i=1}^{n}(x_{1i}-\overline{x_1})^2 = \sum_{i=1}^{n}x_{1i}^2 - n\overline{x_1}^2 = 864585 - 25*34491.92 = 2287.04$$

$$S_{yy} = \sum_{i=1}^{n}(y_{1i}-\overline{y_1})^2 = \sum_{i=1}^{n}y_{1i}^2 - n\overline{y_1}^2 = 847348 - 25*33797.15 = 2419.36$$

$$S_{xy} = \sum_{i=1}^{n}(x_{1i}-\bar{x})(y_{1i}-\bar{y}) = \sum_{i=1}^{n}x_{1i}y_{1i} - n\overline{x_1}\,\overline{y_1} = 855241 - 25*34142.76 = 1671.88$$

$$R_{xy} = \frac{1671.88}{\sqrt{2287.04}\sqrt{2491.36}} = 0.711$$

표본상관계수는 0.711로 1에 가까운 값을 갖는 것으로 나타났으며 이는 첫째 아들과 둘째 아들의 머리 길이 사이에는 양의 관계가 존재한다는 것을 의미한다.

· 머리 폭

$$S_{xx} = \sum_{i=1}^{n}(x_{2i}-\overline{x_2})^2 = \sum_{i=1}^{n}x_{2i}^2 - n\overline{x_2}^2 = 572236 - 25*22837.25 = 1304.64$$

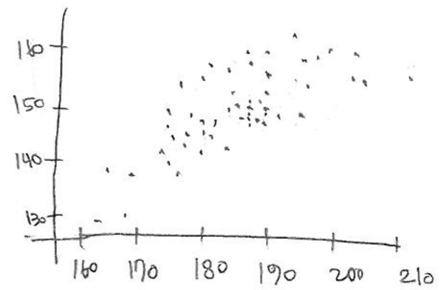
$$S_{yy} = \sum_{i=1}^{n}(y_{2i}-\overline{y_2})^2 = \sum_{i=1}^{n}y_{2i}^2 - n\overline{y_2}^2 = 557895 - 25*22272.58 = 1080.56$$

$$S_{xy} = \sum_{i=1}^{n}(x_{2i}-\bar{x})(y_{2i}-\bar{y}) = \sum_{i=1}^{n}x_{2i}y_{2i} - n\overline{x_2}\,\overline{y_2} = 564670 - 25*22553.15 = 841.28$$

$$R_{xy} = \frac{1080.56}{\sqrt{1304.64}\sqrt{1080.56}} = 0.709$$

표본상관계수는 0.709로 1에 가까운 값을 갖는 것으로 나타났으며 이는 첫째 아들과 둘째 아들의 머리 크기 폭 사이에는 양의 관계가 존재한다는 것을 의미한다.

(3) ①



②

$$r = \frac{S_{(x+z)(y+w)}}{\sqrt{S_{(x+z)(x+z)}}\sqrt{S_{(y+w)(y+w)}}}$$

$$= \frac{702919 + 687264 - (50\times(184.78)(150.18))}{\sqrt{1711933-(184.78^2\times 50)}\sqrt{1130131-(150.18^2\times 50)}}$$

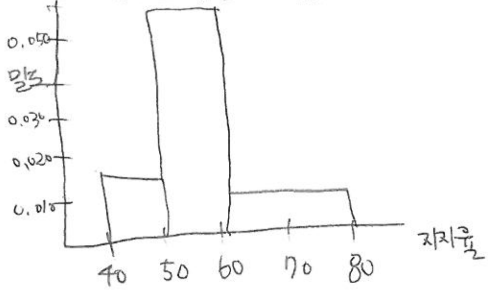
$$= \frac{2669.98}{68.924\times 49.289} = 0.786$$

#2.11

(1) 구간별 밀도를 구하면

구간	밀도	상대도수
(40,50]	0.0176	0.176
(50,60]	0.0588	0.588
(60,80]	0.0118	0.236

이므로 히스토그램을 그리면



이다.

(2) ① 표본평균 : $(49.75+\cdots+77.96)\div 17$

$= 58.216$

표본중앙값 : 9번째 값 $= 56.12$

② 범위: $77.96-49.75 = 28.21$

표본분산 : $\frac{1}{16}\left(\sum x_i^2 - 17\bar{x}^2\right)$

$=\frac{1}{16}(59004.64 - 57614.745)$

$= 86.868.$

표준편차 $\fallingdotseq \sqrt{S^2} = \sqrt{86.868} = 9.320.$

사분위범위 : $Q_3 - Q_1$

Q_1 위치: 5, Q_3 위치: 13

$\therefore Q_1 = 50.43,\ Q_3 = 59.97$

$\therefore 59.97 - 50.43 = 9.54$

변동계수 : $CV = \frac{s}{\bar{x}}\times 100$

$= \frac{9.320}{58.216}\times 100 = 16.009$

- 50 백분위수 $k = 16\times 0.5+1 = 9$

$\therefore 56.12.$

③

- 5 백분위수

$k = 16\times 0.05+1$

$= 1.8$

$49.75\times 0.2 + 49.76\times 0.8 = 49.758$

- 10 백분위수

$k=16\times 0.1+1$

$=2.6$

$49.76\times 0.4 + 49.95\times 0.6 = 49.874$

- 90 백분위수

$k = 16\times 0.9+1 = 15.4$

$69.23\times 0.6 + 77.73\times 0.4 = 72.63$

- 95 백분위수

$k = 16\times 0.95+1 = 16.2$

$77.73\times 0.8 + 77.96\times 0.2 = 77.776.$

④

왜도 $= \sqrt{b_1} = \frac{1}{16}\sum\left(\frac{x-58.216}{9.320}\right)^3$

$= 1.033$

첨도 $b_2 = \frac{1}{16}\sum\left(\frac{x-58.216}{9.320}\right)^4$

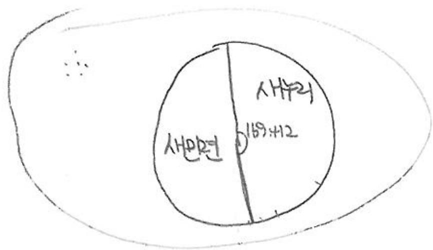
$= 2.833$

(3) 각정당별 비율은

새민련 $\frac{9}{17}$, 새누리 $\frac{8}{17}$ 이므로

파이차트의 각도는 새민련 190.588

새누리 169.412 이다.



(4)

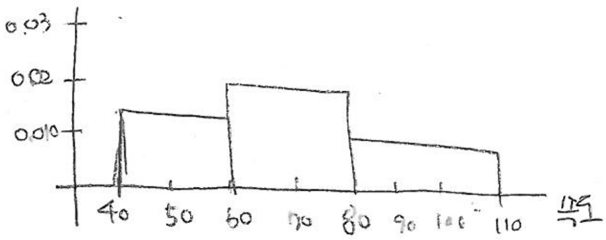
		과반(50)이상 득표 여부	
		o	x
소속정당	새민련	7	2
	새누리	7	1

#2.12 . ㄱ

(1) 구간별 밀도를 구하면

구간	밀도	상대도수
[40, 60)	0.015	0.3
[60, 80)	0.02	0.4
[80, 110)	0.01	0.3

이므로 히스토그램을 그리면



이다.

(2) $(45+51+59+\cdots+102)/10 = 70.1$

(3)

$$\sqrt{\frac{1}{9}\left(\sum x_i^2 - 10\times 70.1^2\right)} = \sqrt{\frac{1}{9}(51851 - 49140.1)}$$

$$= \sqrt{301.2111} = 17.355$$