2장 자료의 정리와 요약

표본 평균 $\overline{x} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n} x_i$

표본 중앙값 $\tilde{x} = \begin{cases} x_{n+1/2}\ (n: odd) \\ \frac{1}{2}\left\{x_{n/2} + x_{n/2+1}\right\}\ (n: even) \end{cases}$

편차 = 관측값(x_i) − 표본평균$(\overline{x})$, $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$

표본분산 $S^2 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - n\overline{x}^2)$

표본 표준 편차 S

표본 범위 $R = \text{가장큰관측값} - \text{가장작은관측값}$

표본사분위 범위$IQR = Q_3 - Q_1$

4장 확률변수

X의 분산 $Var(X) = E(X^2) - \{E(X)\}^2 = \sum_{i=1}^{n}(x_i - \mu)^2 p_i$

X의 표준편차 $\sigma = sd(X) = \sqrt{Var(X)}$

기댓값과 분산의 성질 $E(aX + bY + c) = aE(X) + b(Y) + c, V(aX + b) = a^2V(X)$

이항확률변수의 확률질량함수 $f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}, q = 1 - p$

이항분포의 평균 $X\sim B(n,p), E(X) = np$

이항분포의 표준편차 $X\sim B(n,p), Var(X) = npq$

정규분포의 표준화 $X\sim N(\mu,\sigma^2) \rightarrow Z = \frac{X-\mu}{\sigma}\sim N(0,1)$

정규분포의 확률계산 $\mathrm{P(a \le X \le b) = P}\left(\frac{\mathrm{a}-\mu}{\sigma} \le \mathrm{Z} \le \frac{\mathrm{b}-\mu}{\sigma}\right)$

연속성의 수정을 이용한 근사 확률 계산

$$P(a \le X \le b) \approx P\left(\frac{a-0{,}5-np}{\sqrt{np(1-q)}} < Z < \frac{b+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$P(X = b) \approx P(\frac{b-0{,}5-np}{\sqrt{np(1-q)}} < Z < \frac{b+0{,}5-np}{\sqrt{np(1-q)}})$$

5장

중심 극한 정리

평균 μ, 분산 σ^2인 임의의 모집단에서 표본의 크기 n이 충분히 크면$(n \ge 30)$

표본 평균 $\overline{X}$는 모집단의 분포에 상관없이 근사적으로 정규분포 $N(\mu,\frac{\sigma^2}{n})$을 따른다.

$$\overline{X} \approx N\left(\mu,\frac{\sigma^2}{n}\right), Z = \frac{(\overline{X}-\mu)}{\sigma/\sqrt{n}} \approx N(0,1)$$