9장 단일모집단에 대한 통계적 추론

9.1

(1) $\overline{x} = 17/6 = 2.833$, $s^2 = (101 - 6 \times 2.833^2)/5 = 10.567 \ \Rightarrow\ s = \sqrt{10.567} = 3.251$

(2) μ: 평균 스트레스 증가량

- 가설: $H_0 : \mu \le 0$ vs $H_1 : \mu > 0$
- 검정통계량: $\dfrac{\overline{X}}{S/\sqrt{n}} \sim t_{n-1}$
- $t = \dfrac{2.833 - 0}{3.251/\sqrt{6}} = 2.135 > t_{0.05.5} = 2.015$ 이므로 5% 유의수준에서 귀무가설 기각

  ⇒ 기초통계학이 스트레스를 증가시키는 요인이라고 할 수 있음

(3) $\left[\sqrt{\dfrac{(n-1)S^2}{\chi_{\alpha/2,n-1}}}, \sqrt{\dfrac{(n-1)S^2}{\chi_{1-\alpha/2,n-1}}}\right] \Rightarrow \left[\sqrt{\dfrac{5 \times 10.567}{12.83}}, \sqrt{\dfrac{5 \times 10.567}{0.83}}\right]$ = [2.029, 7.973]

9.2 이 문제는 비올 확률이 60%에 대한 예측력이 적절한지를 알아보기 위한 것임. 비가 내리 비율 $p = 51/100 = 0.51$을 이용하여 비올 확률 θ에 대한 95% 구간추정은

$$p \pm 1.96\sqrt{p(1-p)/100} \ \Rightarrow \ [0.412,\ 0.608]$$

통상적인 신뢰수준 95%에서 구간추정에 0.6이 포함되어 있어 예측력에 큰 문제가 없다고 할 수 있음.

- 비올 확률이 50%라고 할 때, $\left|\dfrac{x/100 - 0.5}{\sqrt{0.5 \times 0.5/100}}\right| = 1.96$ 를 만족하는 x ⇨ $|x - 50| = 9.8$

  ⇒ 40일 이하이거나 60일 이상 비가 내리면 예측력에 문제가 있다고 볼 수 있음

9.3 대표본이므로 정규근사를 이용할 수 있음

- 민간부문: [2232.497, 2299.503]
- 공공부문: [2542.364, 2677.636]
- 노조사업장: [2656.853, 2729.147]
- 비노조사업장: [2409.216, 2488.784]

9.4

(1) $\overline{x}$ = 85.722, s^2 = 292.918

(2) $85.722 \pm 2.11\sqrt{\dfrac{292.918}{18}}$ ⇒ [77.211, 94.233]

(3) $\left[\sqrt{\dfrac{(n-1)S^2}{\chi_{\alpha/2,n-1}}}, \sqrt{\dfrac{(n-1)S^2}{\chi_{1-\alpha/2,n-1}}}\right] \Rightarrow \left[\sqrt{\dfrac{17 \times 292.918}{30.19}}, \sqrt{\dfrac{17 \times 292.918}{7.56}}\right]$ = [12.843, 25.665]

(4) μ: 평균호가

- 가설: $H_0 : \mu \le 70$ vs $H_1 : \mu > 70$
- $t = \dfrac{85.722 - 70}{\sqrt{292.918/18}} = 3.897 > t_{0.05,17} = 1.74$이므로 5% 유의수준에서 귀무가설 기각

  ⇒ 호가가 실제 월세보다 높다고 할 수 있음

9.5

(1) $\bar{x} = 188/10 = 18.8$, $s^2 = 51.2889$ $\Rightarrow$ $s = 7.1616$

(2) $18.8 \pm 2.262 \times 7.1616/\sqrt{10}$ = [13.677, 23.923]

(3) Isoxya cicatricosas라고 하면 $\mu = 28.12$ 이므로 $t = \dfrac{18.8 - 28.12}{7.1616/\sqrt{10}} = -4.115$로 현재자료는 비정상적인 자료라고 할 수 있는 반면 Araneus rufipalpus라고 가정하면 $\mu = 15.66$이고 $t = \dfrac{18.8 - 15.66}{7.1616/\sqrt{10}} = 1.387$이므로 비정상적인 자료라고 볼 수 없음 ⇒ 이 거미는 Araneus rufipalpus 가능성이 더 높음

9.6 $n = 15$, $\bar{x} = 27.8$, $s = 5.5$

(1) $27.8 \pm 2.145 \times 5.5/\sqrt{15}$ = [24.754, 30.846]

(2) 코끼리 수명은 $N(\mu, \sigma^2)$ 따르고 15마리 코끼리는 모집단으로부터 무작위로 선정

(3) μ: 아프리카 코끼리 평균수명

- 가설: $H_0 : \mu \geq 60$ vs $H_1 : \mu < 60$
- $t = \dfrac{27.8 - 60}{5.5/\sqrt{15}} = -23.24 < -2.624 = -t_{0.01,14}$ 이므로 1% 유의수준에서 귀무가설 기각

  ⇒ 아프리카 코끼리 수명은 인도코끼리 수명보다 짧다고 할 수 있음

9.7 유병율: $\theta = 0.248$, 표본유병율: $p = 68/200 = 0.34$, $n = 200$ ⇨ 정규근사가능

(1) 200×0.248 = 49.6(명)

(2) $0.34 \pm 1.96\sqrt{0.34 \times 0.66/200}$ = [0.2743, 0.4057]

(3) 가설: $H_0 : \theta = 0.248$ vs $H_1 : \theta \neq = 0.248$

- $|z| = \left| \dfrac{0.34 - 0.248}{\sqrt{0.248 \times 0.752/200}} \right| = 3.068 > 1.96 = z_{0.025}$이므로 5% 유의수준에서 귀무가설 기각

  ⇒ 고혈압 유병율이 24.8%라는 주장은 문제가 있음

(4) 의료기록은 병원에 방문한 사람, 즉 환자들을 많이 포함되어 있을 가능성이 높아 40대 전체를 대표하는 표본으로 적절하지 않을 수 있음. 단순 치료를 위한 의료기록보다는 건강검진 자료와 같은 자료에서 표본을 선정하는 것이 적절함

9.8

(1) $1.96\sqrt{\dfrac{0.5 \times 0.5}{n}} = 0.015$ $\Rightarrow$ $n = 4268.4$ (4269개)

(2) $n = 200$,

① 문제수정: "만약 θ가 0.3이라면 ~~정상~~ 불량제품"

- $\theta = 0.3$ ⇨ $P(0.27 \leq P \leq 0.33) = P\left(\dfrac{0.27 - 0.3}{\sqrt{0.3 \times 0.7/200}} \leq Z \leq \dfrac{0.33 - 0.3}{\sqrt{0.3 \times 0.7/200}}\right)$
  $= P(-0.93 \leq Z \leq 0.93) = 0.6476$

② 장상제품비율: $174/200 = 0.87$ ⇨ 불량제품비율 $p = 0.13$

- $0.13 \pm 1.645\sqrt{\dfrac{0.13 \times 0.87}{200}}$ = [0.0989, 0.1691]

③ (a) $H_0 : \theta \geq 0.2$ vs $H_1 : \theta < 0.2$

(b) 175개 이상 정상 ⇨ 25개 미만 불량

$\alpha = P_{H_0}(X < 25) = P_{H_0}(X \leq 24.5) \approx P\left(Z \leq \dfrac{24.5-40}{\sqrt{36}}\right) = P(Z \leq -2.5833) = 0.0049$

(c) 179개 정상 ⇨ 21개 불량 ⇒ $p = 21/200 = 0.105$

$z = \dfrac{0.105 - 0.2}{\sqrt{02 \times 0.8/200}} = -3.359 < -1.645$ 이므로 5% 유의수준에서 귀무가설 기각

⇒ 불량률이 20% 미만이라고 할 수 있음

(d) p-값: $P(Z \leq -3.359) = 0.0004$

9.9 문제수정: 2번째 줄 "4.5mg~~이고 표준편차는 0.5mg~~이라고 한다."

$n =$ 100, $\bar{x} =$ 4.6, $s =$ 0.45 ⇨ 대표본으로 정규근사 가능

(1) $4.6 \pm 1.96 \times 0.45/\sqrt{100} =$ [4.5118, 4.6882]

(2) μ: 실제 평균 타르함량

- 가설: $H_0 : \mu \leq 4.5$ vs $H_1 : \mu > 4.5$
- $z = \dfrac{4.6 - 4.5}{0.45/\sqrt{100}} = 2.222 > 1.645 = z_{0.05}$ 이므로 5% 유의수준에서 귀무가설 기각

  ⇒ 실제 평균 타르 함량은 4.5보다 많다고 할 수 있음

9.10 기존 시스템에서의 위험상태에 빠지는 비율: $P(X > 30) = P\left(Z > \dfrac{30-25}{5}\right) = P(Z > 1) = 0.1587$

- 이송시간을 단축시켰다는 것은 위험상태에 빠진 비율이 낮아졌다는 것을 의미
- θ : 위험상태에 빠진 비율
- 가설: $H_0 : \theta \geq 0.1587$ vs $H_1 : \theta < 0.1587$
- $z = \dfrac{5/100 - 0.1587}{\sqrt{0.1587 \times 0.8413/100}} = -2.975 < -1.645$ 이므로 5% 유의수준에서 귀무가설 기각

  ⇒ 새로운 이송시스템은 이송시간을 단축시켰다고 볼 수 있음