문제 **10-1.** 모분산에 대한 추정

**A** 정밀기계 공장에서 제공하는 부품은 표준편차가 **2mm**를 넘지 않도록 요구하고 있다. 만일 표본으로 수집된 데이터의 표준편차가 **2mm**를 넘을 것이라는 뚜렷한 증거가 발생하면, 해당 부품을 모두 폐기 처리하려고 한다. 표본 **20**개를 무작위로 뽑아 조사한 결과 표본표준편차는 $\sqrt{5}$**mm**로 나타났다. 다음 물음에 답하여라.

**1)** 관련된 모수를 기술하여라

관련 모수는 모표준편차 σ **-** 공장 전체 부품 치수 오차의 산포
$\sigma \leq 2mm$

**2)** 모수에 대한 점추정량과 표준오차는**?**

표본 크기: 20
자유도: 20 -1 = 19
표본표준편차 = $\sqrt{5} \approx 2.236mm$
점추정량 = 2.236mm
표준오차 $\frac{2.236}{\sqrt{38}} \approx 0.36mm$

**3)** 모수의 신뢰수준 **95%**에 대한 신뢰구간은**?**

$$\sigma \in \left(\sqrt{\frac{\nu s^2}{\chi^2_{0.975,\nu}}}, \sqrt{\frac{\nu s^2}{\chi^2_{0.025,\nu}}}\right)$$

$$\chi^2_{0.975,19} = 8.907 , \chi^2_{0.025,19} = 32.852$$

$$\sigma_L = \sqrt{\frac{19\times5}{32.852}} = 1.70mm$$

$$\sigma_U = \sqrt{\frac{19\times5}{8.907}} = 3.27mm$$

95% 신뢰 구간 1.70mm < σ < 3.27mm

**4)** 이 데이터를 분석하였을 때**, A** 공장의 부품이 표준편차 **2mm**를 초과한다는 뚜렷한 증거를 확인할 수 있었는가**?**

표본에서 계산한 표준편차는 약 2.24mm이다.
통계적으로 모집단 표준편차가 어디에 있을지 95% 신뢰구간으로 잡아보니 1.70mm ~ 3.27mm 범위가 나왔다.

표본 20개를 놓고 보면, 2mm를 초과한다는 명확한 증거는 없다.

문제 **10-2**
표본크기의 결정 **(**모비율**)**

어느 여론조사기관에서는 정부의 정책지지율을 평가하기 위하여 표본설계를 진행하고 있다**.**

**(1)** **95%** 신뢰수준에서 **1%** 이상 오차가 나타나지 않게 하기 위해서는 어느 정도의 표본을 조사하여야 하는가**?**

p = 0.5로 계산

$n = \frac{1.96^2 \times 0.5 \times 0.5}{0.01^1} = \frac{3.8416 \times 0.25}{0.0001} \approx 9604$
최소 9605명을 조사하면 **+-1%**p 정확도를 보장할 수 있다.

**(2)** 과거 경험에 의하면 국정지지율은 **65%** 정도 나오는 것으로 확인되었다**.** 이러한 정보를 이용하여 표본의 크기를 정하고자 한다면 어느 정도의 표본을 조사하여야 하는가**?**

$p(1-p) = 0.65 \times 0.35 = 0.2275$
$n = \frac{1.96^2 \times 0.2275}{0.01^2} = \frac{3.8416 \times 0.2275}{0.0001} \approx 8744.3$
최소 8745명

문제 **10-3**
**TV** 공장에서는 컬러 **TV**용 튜너의 불량이 과거 데이터에 의해 **5.5%**로 집계되었다**.** 이 불량률을 줄이기 위해 콘덴서의 예비 가열공정을 추가시켰는데**,** 이에 따른 튜너의 불량률이 감소되었는가를 확인하기 위해 새 공정에서 만들어진 **200**개의 튜너를 랜덤하게 채취하여 검사하였더니 **4**개가 불량이었다**.** 이 불량률에 대한 **95%** 신뢰구간을 구해보자**.**

검사한 튜너 수 n = 200
불량 개수 x = 4
표본 불량률 0.02 (2%)

95% 신뢰구간 계산

$Lower = Beta^{-1}(\frac{\alpha}{2}; x, n-x+1),\ Upper = Beta^{-1}(1-\frac{\alpha}{2}; x+1, n-x)$
$(\alpha = 0.05,\ x = 4,\ n = 200)$

$0.5\% < p < 5.2\%$

문제 **10-4**
연못 안에 몇 마리의 물고기가 살고 있는지 알고 싶은 겨울르 생각해봅시다**.** 이것을 어떻게 추정할 수 있을까요**?** 연구자들이 개발한 한 가지 방법은 포획**-**재포획**(capture-recapture)** 방법입니다**.** 이 방법은 개체군의 일부를 포획하고**,** 표식을 한 후 풀어줍니다**.** 나중에 다시 일부를 포획하고**,** 그 안에 표식된 개체가 몇 마리 있는지 세는 것입니다**.**

처음 포획에서는 물고기 **17**마리를 잡아서 모두 표식을 하고 다시 연못에 풀어주었습니다**.** 두 번째 포획에서는 무작위로 물고기 **14**마리를 다시 잡았는데 이 중 표식된 물고기가 **6**마리였습니다**.**

**1)** 연못에 있는 전체 물고기 수가 **25**마리일 확률은 얼마입니까**?**

모수 N에 대한 likelihood를 계산한다.

$$P\left(k\ ||N, M, n\right)\ =\ \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

$$P\left(k\ =\ 6\ ||N\ =\ 25\right)\ =\ \frac{\binom{17}{6}\binom{25-17}{8}}{\binom{25}{14}}$$

$\approx\ 0.278\%$

**2)** 전체 물고기 수로 가장 가능성 높은 값은 얼마입니까**?**

$$\hat{N_{MLE}} =\ \arg max\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

$\hat{N} = 39$

39마리가 전체 개체 수일 때, 관측 결과 k = 6이 나올 likelihood가 가장 높다.