2018 年度秋学期　明治大学理工学部「統計学」期末試験解答例　担当：宮部賢志

10 点 $\times 10$ 問.

期待値 μ, 分散 σ^2 となる正規分布を $N(\mu,\sigma^2)$ で表す.

独立な確率変数 $Z_i,\cdots,Z_k\sim N(0,1)$ の 2 乗和 $Z_1^2+\cdots+Z_k^2$ は自由度 k の χ^2 分布に従い，この分布を χ_k^2 で表す.

$Z\sim N(0,1)$, $U\sim\chi_m^2$, Z,U は独立とする．$T=\frac{Z}{\sqrt{U/m}}$ は自由度 m の t 分布に従い，この分布を t_m で表す．また $t_{m,\alpha}$ を $P(T>t_{m,\alpha})=\alpha$ となる点として定義する.

問題 1. 正規分布 $N(\mu,\sigma^2)$ に従う母集団から，標本 X_1, $\cdots$, X_n を無作為抽出する.

(1) 標本平均 $\overline{X}=\frac{1}{n}\sum_{k=1}^n X_k$ はどんな分布をするか．またこの標本平均の期待値・分散を求めよ.

(2) 不偏分散 $V^2=\frac{1}{n-1}\sum_{k=1}^n(X_k-\overline{X})^2$ が σ^2 の不偏推定量であることを示せ.

(3) $(n-1)V^2/\sigma^2\sim\chi_{n-1}^2$ であることから，V^2 の期待値を求めよ.

(4) $T=\frac{\overline{X}-\mu}{V/c}\sim t_{n-1}$ となるような c を n の式で書け.

(5) μ の 95% 信頼区間を，n, $\overline{X}$, V, $T_{m,\alpha}$ を使って表わせ.

解答.　(1) $E[\overline{X}]=\mu$, $V[\overline{X}]=\frac{\sigma^2}{n}$ である．また正規分布の再帰性より $\overline{X}$ も正規分布する.

(2)

$$\begin{aligned}E[\sum_{k=1}^n(X_k-\overline{X})^2]=&E[\sum_{k=1}^n((X_k-\mu)-(\overline{X}-\mu))^2]\\=&\sum_{k=1}^n E[(X_k-\mu)^2]-2\sum_{k=1}^n E[(X_k-\mu)\cdot\frac{X_k-\mu}{n}]+nE[(\overline{X}-\mu)^2]\\=&n\sigma^2-2\sigma^2+\sigma^2=(n-1)\sigma^2\end{aligned}$$

より $E[V^2]=\sigma^2$ である.

(3) 自由度 k の χ^2 分布の期待値は k であるから，$E[(n-1)V^2/\sigma^2]=n-1$ より，$E[V^2]=\sigma^2$.

(4) $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$, $(n-1)V^2/\sigma^2\sim\chi_{n-1}^2$ より，$T=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\times\frac{\sqrt{n-1}}{\sqrt{(n-1)V^2/\sigma^2}}\sim t_{n-1}$. すなわち $c=\sqrt{n}$.

(5) $\left|\frac{\overline{X}-\mu}{V/\sqrt{n}}\right|\le t_{n-1,0.025}$ を μ に関して解いて，信頼区間は $\overline{X}\pm t_{n-1,0.025}\frac{V}{\sqrt{n}}$.

□

X の同時確率関数 f もしくは同時確率密度関数 f において x を X に置き換えたものを $L(\theta|X)=\prod_{i=1}^n f(X_i|\theta)$ を尤度関数という．尤度関数を最大とする $\hat{\theta}$ を最尤推定量という．因子分解定理より，$f(x|\theta)=h(x)g(T(x)|\theta)$ と書けるなら，$T(X)$ は十分統計量である.

X の θ に対するフィッシャー情報量 $I(\theta)$ は，$I(\theta)=E[(\frac{d}{d\theta}\log f(X|\theta))^2]=-E[\frac{d^2}{d\theta^2}\log f(X|\theta)]$ で計算できる．クラメール・ラオの不等式より，不偏推定量 $\hat{\theta}$ の分散の下限は $\frac{1}{I(\theta)}$ である.

問題 2. $X_1,\cdots,X_n$ は独立で，$P(X_i=1)=p\in(0,1)$, $P(X_i=0)=1-p$ とする．$\hat{p}=\overline{X}=\frac{1}{n}\sum_{k=1}^n X_k$ とする.

(1) $\hat{p}$ は p の不偏推定量であることを示せ.

(2) $\hat{p}$ は最尤推定量であることを示せ.

(3) $\hat{p}$ は十分統計量であることを示せ．

(4) $\hat{p}$ の分散はクラメール・ラオの下限と一致することを示せ．

(5) $\hat{p}$ は p の最良線形不偏推定量 (BLUE) であることを示せ．

解答． (1) $E[X_k] = p$ であるから，$E[\hat{p}] = \frac{1}{n} \cdot nE[X_1] = p$.

(2) $S = \sum_{k=1}^n X_k$ とおく．

$$L(p|X) = p^S (1-p)^{n-S}$$

より，

$$\log L(p|X) = S \log p + (n-S) \log(1-p)$$

を最大化すれば良い．導関数 $\frac{S}{p} - \frac{n-S}{1-p} = 0$ より $(1-p)S = (n-S)p$，$p = \frac{S}{n}$ で最大値をとる．これは $\hat{p}$ が最尤推定量であることを意味する．

(3) $s = \sum_{k=1}^n x_k$ とおくと，$f(x|p) = p^s (1-p)^{n-s}$ であり，因子分解定理から S や $\hat{p}$ が十分統計量であることが分かる．

(4)

$$\log f(X|p) = S \log p + (n-S) \log(1-p)$$

より

$$\begin{aligned} \frac{d}{dp} \log f(X|p) =& \frac{S}{p} - \frac{n-S}{1-p} \\ \frac{d^2}{dp^2} \log f(X|p) =& -\frac{S}{p^2} - \frac{n-S}{(1-p)^2} \end{aligned}$$

よって，

$$I(p) = \frac{np}{p^2} + \frac{n-np}{(1-p)^2} = \frac{n}{p} + \frac{n}{1-p} = \frac{n}{p(1-p)}$$

一方，

$$V[\hat{p}] = E[(\overline{X} - p)^2] = \frac{1}{n^2} \sum_{k=1}^n E[(X_k - p)^2] = \frac{p(1-p)}{n}$$

これは $\frac{1}{I(p)}$ である．

(5) $\hat{p}$ は線形で，p の不偏推定で，分散は最小となることから，BLUE である．

□