RSS Higher Certificate in Statistics, Specimen A

Module 5 : Further Probability and Inference

Solutions

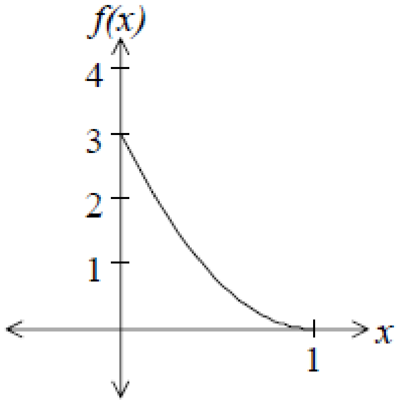
Question 1

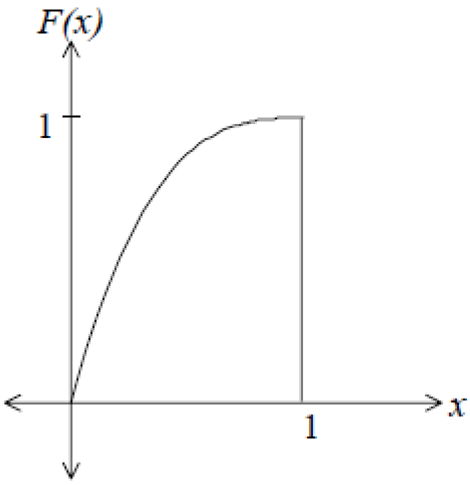
$f(x) = \alpha(1-x)^{\alpha-1}, \quad 0 < x < 1, \quad \alpha > 0$

(i) $F(x) = \int_0^x \alpha(1-u)^{\alpha-1}\,du = \left[-(1-u)^{\alpha}\right]_0^x = 1-(1-x)^{\alpha}$ (for $0 < x < 1$ and $\alpha > 0$) .

中央値 m は $F(m) = 1/2$ で与えられるので, $1 - (1 - m)^{\alpha} = 1/2$, $(1 - m)^{\alpha} = 1/2$, $1 - m = 2^{-1/\alpha}$.
よって $m = 1 - 2^{-1/\alpha}$.

$\alpha = 3$ のとき, $f(x) = 3(1 - x)^2$, $F(x) = 1 - (1 - x)^3$ ($0 < x < 1$) であるので概形は以下のようになる.





(ii) 尤度関数は $L = \prod_{i=1}^{n} \alpha(1-x_i)^{\alpha-1} = \alpha^n \prod_{i=1}^{n}(1-x_i)^{\alpha-1}$ であり, 対数尤度関数は

$\log L = n\log\alpha + (\alpha-1)\sum_{i=1}^{n}\log(1-x_i)$ である. これを α について微分すると

$\dfrac{d\log L}{d\alpha} = \dfrac{n}{\alpha} + \sum_{i=1}^{n}\log(1-x_i)$ となるので, これを 0 とおくことにより最尤推定量

$\hat{\alpha} = \dfrac{-n}{\sum_{i=1}^{n}\log(1-x_i)}$ が得られる.

($\dfrac{d^2\log L}{d\alpha^2}$ を考えると, これが最大値を与えることが確かめられる.)

$\dfrac{d^2 \log L}{d\alpha^2} = -\dfrac{n}{\alpha^2}$ であるから，問題文に与えられた仮定を用いると $\hat{\alpha}$ は近似的に平均 α ，分散 α^2/n の正規分布に従う．分散を $\hat{\alpha}^2/n$ で推定すれば近似的に $\hat{\alpha} \sim \mathrm{N}(\alpha, \hat{\alpha}^2/n)$ となる．

したがって，近似的な 90%信頼区間は

$$0.90 \approx P\left(-1.645 < \frac{\hat{\alpha}-\alpha}{\hat{\alpha}/\sqrt{n}} < 1.645\right)$$

でを整理することにより，区間 $\left(\hat{\alpha} - \dfrac{1.645\hat{\alpha}}{\sqrt{n}}, \hat{\alpha} + \dfrac{1.645\hat{\alpha}}{\sqrt{n}}\right)$ として導かれる．

与えられた観測値によると $n = 5$ で $1-x_i$ の値は 0.88, 0.57, 0.93, 0.13, 0.71 である．このとき

$$\sum \log(1-x_i) = -0.1278 - 0.5621 - 0.0726 - 2.0402 - 0.3425 = -3.1452$$

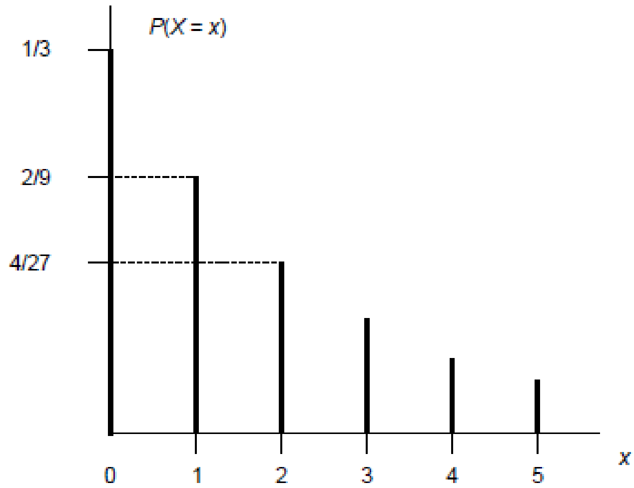
であるので，最尤推定値は $\hat{\alpha} = \dfrac{5}{3.1452} = 1.5897$ となる．

また，$\dfrac{1.645\hat{\alpha}}{\sqrt{n}} = \dfrac{1.645 \times 1.5897}{\sqrt{5}} = 1.1695$ より，90%信頼区間は (0.420, 2.759) となる．

Question 2

確率関数は $f(x) = (1-p)^x p \quad (x = 0, 1, 2, \ldots ; 0 < p < 1)$ である.

(i) $f(x)$ の概形は以下のとおり.



(ii) 確率母関数 $G(s)$は

$$G(s) = E\left[s^X\right] = \sum_{x=0}^{\infty} s^x (1-p)^x p = p\sum_{x=0}^{\infty}\{(1-p)s\}^x = \frac{p}{1-(1-p)s}$$

となる（収束するためには $|s| < 1/(1-p)$ が必要）.

期待値は$E[X] = G'(1)$で与えられる. $G'(s) = p\dfrac{1-p}{\{1-(1-p)s\}^2}$ に $s = 1$ を代入して

$G'(1) = \dfrac{p(1-p)}{p^2}$, すなわち期待値は $\dfrac{1-p}{p}$ である.

分散は $\mathrm{Var}(X) = G''(1) + E[X] - E[X]^2$ で与えられる. $G''(s) = \dfrac{2p(1-p)^2}{\{1-(1-p)s\}^3}$ であるから

$G''(1) = \dfrac{2(1-p)^2}{p^2}$, よって分散は $\dfrac{2(1-p)^2}{p^2} + \dfrac{1-p}{p} - \left(\dfrac{1-p}{p}\right)^2 = \dfrac{1}{p^2}\left\{(1-p)^2 + p(1-p)\right\} = \dfrac{1-p}{p^2}$

となる.

(iii) $U = Y + Z$ とおく. U が値 r $(r = 0, 1, 2, 3, \ldots)$ をとるとき,

$Y = 0, \quad Z = r$
or $Y = 1, \quad Z = r-1$
or $Y = 2, \quad Z = r-2$
…
or $Y = r-1, \quad Z = 1$
or $Y = r, \quad Z = 0 \qquad (*)$

が考えられる.

$p_i = P(Y = i), \pi_i = P(Z = i)$ とすれば, 上の (*) より,

$$P(U = r) = p_0\pi_r + p_1\pi_{r-1} + p_2\pi_{r-2} + \cdots + p_{r-1}\pi_1 + p_r\pi_0$$

となる．

いま，

$$G_Y = E(s^Y) = p_0 s^0 + p_1 s^1 + p_2 s^2 + \cdots + p_{r-1} s^{r-1} + p_r s^r + \cdots,$$

$$G_Z = E(s^Z) = \pi_0 s^0 + \pi_1 s^1 + \pi_2 s^2 + \cdots + \pi_{r-1} s^{r-1} + \pi_r s^r + \cdots$$

であるから，$G_Y G_Z$ の s^r の係数は $p_0\pi_r + p_1\pi_{r-1} + p_2\pi_{r-2} + \cdots + p_{r-1}\pi_1 + p_r\pi_0$ であり，これは $E(s^U)$ の s^r の係数（すなわち $P(U = r)$ ）と等しい．したがって $G_{Y+Z} = G_Y G_Z$ が成り立つ．

（**別解**：Y と Z が独立であることから，$G_{Y+Z} = E(s^{Y+Z}) = E(S^Y)\,E(S^Z) = G_Y G_Z$.）

独立な確率変数 X_1，X_2 は与えられた幾何分布に従うので，直ちに

$$G_{X_1+X_2} = \left(G_X\right)^2 = \frac{p^2}{\left(1-(1-p)s\right)^2}$$

が得られる．

Question 3

(i) 確率変数 X の確率関数または確率密度関数が（一つの）パラメータ θ に依存すると仮定する．推定量 T は無作為抽出された標本値の組 $X_1, X_2, \ldots, X_n$ の関数である．（同じくサイズ n の）標本が何組か得られたら，T は標本ごとに異なる値をとるであろう．この T の「標本分布」や，もしくは少なくともその期待値と分散はしばしば求められ，これらは推定量 T の性質を調べるためによく用いられる．

(A) 「不偏」推定量とは標本分布の平均が θ であるような推定量のことである．すなわち，繰り返し抽出したときに $E(T)=\theta$ となる推定量である．たとえば，標本平均は母平均の不偏推定量である．多くの場合，不偏推定量は非常に役に立つ．なぜならば，抽出した標本が 1 つのみのとき（推定量の値が真の θ の近くにあるという保証はないが－以下の(B) (C)を見よ），この推定量のとる値は，θ の真の値の見当として適切であり理解しやすいからである．いくつかの問題では，単純な変換をほどこすことにより不偏推定量に直せるような偏りを持つ推定量を求めることの方が，より簡単でより自然な場合がある．例として，平均が未知のときの母分散 σ^2 の推定があげられる．偏差平方和を n で割ると不偏でない推定量となるが，n ではなく $n-1$ で割るとバイアスを除くことができる．

(B) バイアスと同様に精度は重要である．標本サイズが大きいほど推定量の精度は高いはずである．なぜならば大きい標本ほどより多くの情報を持っているはずだからである．標本サイズ→∞のときに推定量とパラメータとの差が微小量 ε 以上となる確率が 0 に近づくならば，推定量は「一致性」を持つ．しかし，この基準よりも，「標本サイズ→∞のとき標本分布の分散→ 0 となるならば，推定量は一致性を持つ」という分散に基づく基準の方が使いやすい．（推定量が誤ったところに「収束する」場合，偏りのある推定量に対してこの基準を使うには注意が必要である．標本サイズ→∞のときバイアス自体→0 となるならば，この基準は十分である．）通常の推定量のほぼ全てがこの意味で「一致性」を持つ．たとえば母平均の推定において，標本平均は不偏推定量であり，その分散は σ^2/n であるので，標本平均は母平均の一致推定量である．これが言えない例として Cauchy 分布（自由度 1 の t 分布）があげられる．これは「すその重い」分布であり，すその“重さ”のために平均と分散は（どちらとも）存在しない．しかし，中央値や最頻値は分布の「中央 (centre)」を表しており，標本平均でこれらを推定することは自然であると考えるかもしれない．しかしこの分布では，標本平均の確率密度関数は 1 回の観測値の確率密度関数と同じであり，それゆえ，標本サイズを増やすことのメリットはなく，この場合，標本平均は「中央 (centre)」の一致推定量ではない．

(C) 不偏性と一致性の両方が仮定されている 2 つの推定量を同じ目的で選ぶなら，分散の

小さい方が好ましい．やや乱暴に言えば，分散の小さい推定量の方が，パラメータ θ の真の値に ”より近い” 値をとる可能性が ”より高い”．分散が小さい方の推定量はより「有効」であると言える．たとえば正規分布の母平均 μ の推定において，標本平均の代わりに標本中央値が推定値として扱われるかもしれない．標本平均と標本中央値はどちらも不偏であり一致性を持つ．しかしサンプルサイズが n のときの標本中央値分散は $\pi\sigma^2/2n$ であり，これは標本平均の分散よりも大きいので，有効性に劣る．相対的な有効性は 2 つの推定量の分散の比で測られる．（実際，標本平均がこの推定問題において考えられうる最小の分散を持つので，十分有効である．それに対し，標本中央値は十分に有効ではない（これらの概念は the Graduate Diploma examination の Statistical Theory and Methods で詳しく扱う．そこでは平均 2 乗誤差を用いた偏りのある推定量の拡張について学ぶ）．

(ii) X の確率密度関数は $g(x)=\begin{cases}1/\theta & (0\le x\le\theta)\\ 0 & (\text{otherwise})\end{cases}$ である．

(A) 尤度関数は $(1/\theta)^n=\theta^{-n}$ であり，これはグラフを考えると θ の減少関数である．したがって θ の値が一番小さいとき θ^{-n} は最大となるが，θ が観測値での最大の値よりも小さくなることはないので，尤度が最大となるのは $X_{\max}$ のときである．

(B) $E(\hat{\theta})=E(X_{\max})=\int_0^\theta xf(x)dx=\int_0^\theta \frac{nx^n}{\theta^n}dx=\frac{n\theta}{n+1}$． よって最尤推定量は不偏ではない．

修正された推定量 $\frac{n+1}{n}X_{\max}$ は不偏であることは直ちにわかる．

(C) モーメント法による推定値は 2.6（これは「取りえない値」であることに注意！）であり，修正された最尤推定値は $(7/6)\times 2.8=3.27$ である．

不偏性

両推定量が不偏であることはすでに示した．

一致性

モーメント法による推定量の分散は問題文で $\theta^2/3n$ と与えられており，$n\to\infty$ のとき 0 に収束するので，この推定量は一致性を持つ（これは不偏であるので，分散に基づく基準が直接使えることに注意せよ）．

修正された最尤推定量の分散も問題文に与えられていて，$\theta^2/\{n(n+2)\}$ である．$n\to\infty$ のとき，これは 0 に収束するので，これも一致推定量である．（これは最尤推定量自体が一致性をもつケースであることに注意せよ．$n\to\infty$ のときバイアス $\to$ 0 となるのと同様に分散 $\to$ 0 となる．）

有効性

$$\frac{(\text{モーメント法による推定量の分散})}{(\text{修正された最尤推定量の分散})} = \frac{\theta^2/(3n)}{\theta^2/\{n(n+2)\}} = \frac{n+2}{3}$$

であり，全ての n (>1) についてこの値は 1 より大きい．したがって修正された推定量はモーメント法による推定量より有効である．

Question 4

(i) 表の全 12 項目の和は 30c である．これらの確率は足して 1 なので，c=1/30 である．

(ii) 周辺分布は行や列の合計から与えられる．

よって $P(X = 1) = 15c = 1/2$, $P(X = 2) = 10c = 1/3$, $P(X = 3) = 5c = 1/6$.

同様に $P(Y = 1) = 12c = 2/5$, $P(Y = 2) = 6c = 1/5$, $P(Y = 3) = 6c = 1/5$, $P(Y = 4) = 6c = 1/5$.

(iii) $E(X) = 1\times\frac{1}{2} + 2\times\frac{1}{3} + 3\times\frac{1}{6} = \frac{1}{2} + \frac{2}{3} + \frac{1}{2} = \frac{5}{3}$.

また $E(X^2) = 1\times\frac{1}{2} + 4\times\frac{1}{3} + 9\times\frac{1}{6} = \frac{1}{2} + \frac{4}{3} + \frac{3}{2} = \frac{10}{3}$ であるから

$\mathrm{Var}(X) = \frac{10}{3} - \left(\frac{5}{3}\right)^2 = \frac{5}{9}$ が得られる．

後で $E(Y)$ が必要となるので同様に計算しておくと，$E(Y) = \frac{2}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5} = \frac{11}{5}$ である．

XY の分布は

Values of xy	1	2	3	4	6	12
Probability	$6c$	$7c$	$4c$	$6c$	$5c$	$2c$

(c=1/30)

であるから，

$$E(XY) = 1\times\frac{6}{30} + 2\times\frac{7}{30} + 3\times\frac{4}{30} + 4\times\frac{6}{30} + 6\times\frac{5}{30} + 12\times\frac{2}{30} = \frac{11}{3}$$

であり，また $E(X)E(Y) = \frac{5}{3}\times\frac{11}{5} = \frac{11}{3}$ である．

したがって共分散は $\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$ である．

(iv) $\mathrm{Cov}(X, Y) = 0$ であるにもかかわらず，また $P(X = x, Y = y) = P(X = x)P(Y = y)$を満たす組 (x, y) のセルがあるにもかかわらず X と Y は独立ではない．たとえば，$P(X = 1, Y = 4) = 2/15$ だが $P(X = 1)P(Y = 4) = 1/10$ であるから $P(X = 1, Y = 4) \neq P(X = 1)P(Y = 4)$ である．

(v) $U = 1$ ($X = 1$ または $X = 3$ のとき), $U = 0$ ($X = 2$ のとき), $V = 1$ ($Y = 1$ または $Y = 3$ のとき), $V = 1$ ($Y = 2$ または $Y = 4$ のとき) であるから U と V の同時分布と周辺分布は下の表のようになる．

		Values of V 0	1	
Values of U	0	$2c$ = 1/15	$8c$ = 4/15	$10c$ = 1/3
	1	$10c$ = 1/3	$10c$ = 1/3	$20c$ = 2/3
		$12c$ = 2/5	$18c$ = 3/5	

たとえば $(U, V) = (0, 0)$ のセルを考える．このセルの確率は 1/15 だが周辺確率の積は 2/15 である．したがって U と V は独立でない．