수리통계 1. 3차 중간시험 2016. 07. 18

[1] (15점 : (a)5점 (b)5점 (c)5점)
확률변수 X가 구간 $(-1,2)$에서의 균등분포 $U(-1,2)$를 따를 때 다음 각 경우에 Y의 확률밀도함수를 구하여라.

(a) $Y = 2 + 3\log\frac{1+X}{2-X}$ (b) $Y = 2(-\log\frac{X+1}{3})^{1/2}$ (c) $Y = X^2$

[2](20점: (a)10점 (b)10점)
발생률이 λ인 포아송과정 $\{N_t : t \geq 0\}$에서 r번째 현상이 발생할 때까지의 시간을

$$W_r = \min\{t : N_t \geq r\}\ (r = 1, 2, \cdots)$$

이라고 할 때, 다음에 답하여라.

(a) $X = W_2/W_3,\ Y = W_2/W_4$ 의 공분산 $\mathrm{Cov}(X,Y)$를 구하여라.

(b) $X = W_3/W_4,\ Y = W_3/W_5$ 의 결합확률밀도함수를 구하여라.

[3](10점: (a)5점 (b)5점)
확률변수 X_1, X_2가 서로 독립이고 표준정규분포 $N(0,1)$을 따를 때, 다음 확률변수의 확률밀도함수를 구하여라.

(a) $Y = \frac{X_2^4}{X_1^4}$ (b) $Z = \frac{X_2}{\sqrt{X_1^2 + X_2^2}}$

[4] (20점 : (a)10점 (b)10점)
확률밀도함수가

$$f(x) = 3x^2 I_{(0,1)}(x)$$

인 분포로부터의 랜덤표본 $X_1, \cdots, X_n$에 기초한 순서통계량을 $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$이라고 하고

$$Y_1 = -r\log(X_{(r)}/X_{(r+1)})\ (r = 1, \cdots, n),\ X_{(n+1)} \equiv 1$$

이라고 할 때 다음에 답하여라.

(a) $Y_1, Y_2, \cdots, Y_n$ 의 결합확률밀도함수를 구하여라.

(b) $Z = Y_1 + \cdots + Y_k\ (1 \leq k < n)$의 확률밀도함수를 구하여라.

[5] (15점 : (a) 10점 (b) 5점)

$\mathrm{X} = (\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \mathrm{X}_4)' \sim \mathrm{N}(\mu, \Sigma)$이고

$$\mu = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \ \Sigma = \begin{pmatrix} 1 & -1 & 1 & 0 \\ -1 & 2 & 0 & 2 \\ 1 & 0 & 4 & 4 \\ 0 & 2 & 4 & 8 \end{pmatrix}$$

일 때, 다음에 답하여라.

(a) $\mathrm{X}_1 = x_1, \mathrm{X}_2 = x_2$ 가 주어진 경우에 $(\mathrm{X}_3, \mathrm{X}_4)'$의 조건부 분포를 구하여라.

(b) 다음 이차형식의 확률밀도함수를 구하여라.

$\mathrm{Q} = 2(\mathrm{X}_3 - 2\mathrm{X}_1 - \mathrm{X}_2 + 1)^2 + (\mathrm{X}_4 - 2\mathrm{X}_1 - 2\mathrm{X}_2 + 2)^2 - (2\mathrm{X}_3 - 4\mathrm{X}_1 - 2\mathrm{X}_2 + 2)(\mathrm{X}_4 - 2\mathrm{X}_1 - 2\mathrm{X}_2 + 2)$

[6](20점: (a)10점 (b)5점 (c)5점)

$(\mathrm{X}_1, \mathrm{Y}_1)', \cdots, (\mathrm{X}_n, \mathrm{Y}_n)'\ (n > 2)$이 이변량 정규분포

$$N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right), \ \sigma_1^2 > 0, \ \sigma_2^2 > 0, \ -1 < \rho < 1$$

에서의 랜덤표본이고, r이 표본상관계수 즉

$$r = \frac{\sum_1^n (\mathrm{X}_i - \overline{\mathrm{X}})(\mathrm{Y}_i - \overline{\mathrm{Y}})}{\sqrt{\sum_1^n (\mathrm{X}_i - \overline{\mathrm{X}})^2}\sqrt{\sum_1^n (\mathrm{Y}_i - \overline{\mathrm{Y}})^2}}$$

이라고 할 때, 다음에 답하여라.

(a) 표본상관계수의 표본분포에 대하여 다음의 대의적 정의가 성립하는 것을 밝혀라.

$$\frac{r}{\sqrt{1 - r^2}} \stackrel{d}{\equiv} \frac{\sqrt{\mathrm{V}_1}\,\rho / \sqrt{1 - \rho^2} + \mathrm{U}}{\sqrt{\mathrm{V}_2}}$$

여기에서 $\mathrm{V}_1 \sim \chi^2(n-1), \mathrm{V}_2 \sim \chi^2(n-2), \mathrm{U} \sim \mathrm{N}(0,1)$ 이고 $\mathrm{V}_1, \mathrm{V}_2, \mathrm{U}$는 서로 독립이다.

(b) $n > 4$일 때, $\mathrm{Var}\left(\frac{r}{\sqrt{1 - r^2}}\right)$을 구하여라.

(c) q가 주어진 값일 때, 다음 확률이 모상관계수 ρ의 증가함수임을 밝혀라.

$$\mathrm{P}_\rho\{r \geq q\}$$

수리통계 1. 3차 중간시험 2017. 07. 17

[1](15점: (a)5점, (b)5점, (c)5점)
발생률이 λ인 포아송과정 $\{N_t : t \geq 0\}$에서 r번째 현상이 발생할 때까지의 시간을

$$W_r = \min\{t : N_t \geq r\} \ (r = 1, 2, \cdots)$$

이라고 할 때, 다음 각 경우에 X의 확률밀도함수를 구하여라.

(a) $X = 3(-\log(W_1/W_2))^2$ (b) $X = 3 + 2\log\frac{W_1}{W_2 - W_1}$ (c) $X = (3\frac{W_1}{W_2} - 1)^2$

[2](20점: (a)10점 (b)10점)
확률변수 X_1, X_2, X_3가 서로 독립이고 표준정규분포 $N(0,1)$을 따를 때, 다음 확률변수의 확률밀도함수를 구하여라.

(a) $Y = (X_1 + X_2 - \sqrt{2}X_3)/(X_1 + X_2 + \sqrt{2}X_3)$ (b) $Z = \frac{X_3}{\sqrt{X_1^2 + X_2^2 + X_3^2}}$

[3] (15점 : (a)10점, (b)5점)
확률밀도함수가

$$f(x) = \alpha x^{\alpha - 1}\exp(-x^{\alpha})I_{(0,+\infty)}(x) \quad (\alpha > 0)$$

로 주어지는 와이불분포 Weibull$(\alpha, 1)$로 부터의 크기가 $n(n \geq 2)$인 랜덤표본에 기초한 순서통계량을 $X_{(1)} < \cdots < X_{(n)}$이라고 할 때, 다음에 답하여라.

(a) $Y = X_{(n)}/X_{(1)}$의 확률밀도함수가 다음과 같이 주어지는 것을 밝혀라.

$$pdf_Y(y) = \sum_{k=0}^{n-2} n(n-1)\binom{n-2}{k}(-1)^{n-2-k}((k+1) + (n-k-1)y^{\alpha})^{-2}\alpha y^{\alpha-1}I_{(1<y)}$$

(b) $E[\log(X_{(n)}/X_{(1)})]$을 구하여라.

[4] (15점: (a)10점 (b)5점)
확률밀도함수가

$$f(x;\alpha,\beta) = \beta\alpha^{\beta}x^{-\beta-1}I_{[\alpha,\infty)}(x)$$

로 주어지는 두 개의 모수를 갖는 파레토분포 Pareto(α,β), $\alpha > 0, \beta > 0$ 모형에서의 랜덤표본 $X_1, X_2, \cdots, X_n$ $(n \geq 2)$에 기초한 순서통계량을 $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$이라고 할 때, 다음에 답하여라.

(a) $Y = 2\beta\sum_{r=1}^{n}\log(X_r/X_{(1)})$의 확률밀도함수를 구하여라.

(b) $\hat{\beta} = (\frac{1}{n}\sum_{r=1}^{n}\log(X_r/X_{(1)}))^{-1}$을 이용하여 β에 대한 $100(1-\alpha)\%$신뢰구간을 제시하라.

[5] (15점: (a)5점 (b)10점)
선형 회귀 모형

$$\begin{cases} \mathrm{Y} = X\beta + e \\ e \sim \mathrm{N}_n(0, \sigma^2 V) \end{cases}$$

에서, V는 알려져 있는 행렬로서 양의 정부호(positive definite) 행렬이고 주어진 상수의 행렬 X는 $n \times (p+1)$ 행렬로서 계수(rank)가 $(p+1)$이다. 이러한 모형에서는

$$\hat{\beta} = (X^t V^{-1} X)^{-1} X^t V^{-1} \mathrm{Y}$$

로 정의되는 표본회귀계수와 σ^2의 추측값으로는

$$\widehat{\sigma^2} = (\mathrm{Y} - X\hat{\beta})^t V^{-1} (\mathrm{Y} - X\hat{\beta}) / (n - p - 1)$$

을 사용하여 추론을 하게 된다. 이 때 다음에 답하여라.

(a) $\hat{\beta}$ 와 $\widehat{\sigma^2}$ 이 서로 독립임을 밝혀라.

(b) $\frac{1}{p+1}(\hat{\beta} - \beta)^t X^t V^{-1} X (\hat{\beta} - \beta) / \widehat{\sigma^2}$의 분포를 구하고, 이를 이용하여 β의 신뢰집합을 형성하는 방법을 설명하여라.

[6](20점: (a)10점 (b)10점)

$(\mathrm{X}_1, \mathrm{Y}_1)', \cdots, (\mathrm{X}_n, \mathrm{Y}_n)'$ $(n > 2)$이 이변량 정규분포

$$N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right),\ \sigma_1^2 > 0,\ \sigma_2^2 > 0,\ -1 < \rho < 1$$

에서의 랜덤표본이고, r이 표본상관계수 즉

$$r = \frac{\sum_1^n (\mathrm{X}_i - \overline{\mathrm{X}})(\mathrm{Y}_i - \overline{\mathrm{Y}})}{\sqrt{\sum_1^n (\mathrm{X}_i - \overline{\mathrm{X}})^2}\sqrt{\sum_1^n (\mathrm{Y}_i - \overline{\mathrm{Y}})^2}}$$

이라고 할 때, 다음에 답하여라.

(a) 표본상관계수의 표본분포에 대하여 다음의 대의적 정의가 성립하는 것을 밝혀라.

$$\frac{r}{\sqrt{1 - r^2}} \stackrel{d}{\equiv} \frac{\sqrt{\mathrm{V}_1}\,\rho / \sqrt{1 - \rho^2} + \mathrm{Z}}{\sqrt{\mathrm{V}_2}}$$

여기에서 $\mathrm{V}_1 \sim \chi^2(\mathrm{n}-1), \mathrm{V}_2 \sim \chi^2(\mathrm{n}-2), \mathrm{Z} \sim \mathrm{N}(0,1)$ 이고 $\mathrm{V}_1, \mathrm{V}_2, \mathrm{Z}$는 서로 독립이다.

(b) 다음의 확률이 $|\rho|$의 증가함수임을 밝히고, 이 확률의 최소값을 구하여라.

$$\mathrm{P}_\rho\left\{\sqrt{n-2}\,\frac{|r|}{\sqrt{1-r^2}} \geq t_{0.025}(n-2)\right\}$$

수리통계 1. 3차 중간시험 2018. 07. 20

[1](15점: (a)5점, (b)5점, (c)5점)

확률변수 X,Y가 서로 독립이고 표준정규분포 $N(0,1)$을 따를 때, 다음 각 경우에 Z의 확률밀도함수를 구하여라.

(a) $Z = \dfrac{2X+3Y}{X}$ (b) $Z = \dfrac{X^2 - Y^2}{X^2 + Y^2}$ (c) $Z = \dfrac{e^{-(X^2+Y^2)/2}}{1+e^{-(X^2+Y^2)/2}}$

[2] (15점)

$U_{(1)} < U_{(2)} < \cdots < U_{(n)}$이 균등분포 $U(0,1)$로부터의 랜덤표본 $U_1, U_2, \cdots, U_n$에 기초 한 순서통계량일 때,

$$Y_1 = \frac{U_{(1)}}{U_{(2)}}, Y_2 = \frac{U_{(2)}}{U_{(3)}}, \cdots, Y_{n-1} = \frac{U_{(n-1)}}{U_{(n)}}, Y_n = U_{(n)}$$

의 결합확률밀도함수를 구하여라.

[3] (15점)

이차원 확률변수 $(X,Y)^t$의 결합분포는 이변량정규분포가 아니면서, X와 Y 각각의 주변분포는 표준정규분포 $N(0,1)$인 예를 들고. 그 이유를 밝혀라

[4] (20점: (a)10점 (b)10점)

확률밀도함수가

$$f(x;\mu,\sigma) = \frac{1}{\sigma} e^{(x-\mu)/\sigma} I_{[\mu,\infty)}(x)$$

로 주어지는 두 개의 모수를 갖는 지수분포 $\mathrm{Exp}(\mu,\sigma)$, $-\infty < \mu < +\infty, \sigma > 0$ 모형에서의 랜덤표본 $X_1, X_2, \cdots, X_n$ $(n \geq 2)$에 기초한 순서통계량을 $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$이라고 할 때,
다음에 답하여라.

(a) $Y = \sum_{i=1}^{n} (X_i - X_{(1)})$의 확률밀도함수를 구하여라.

(b) $\hat{\mu} = X_{(1)}, \hat{\sigma} = \sum_{i=1}^{n} (X_i - X_{(1)})/(n-1)$이라고 할 때,

$$\frac{n(\hat{\mu} - \mu)}{\hat{\sigma}}$$

의 분포를 이용하여 μ에 대한 $100(1-\alpha)\%$신뢰구간을 제시하라.

[5] (15점:(a)5점, (b)5점, (c)5점)

선형 회귀 모형

$$\begin{cases} \mathrm{Y} = X\beta + e \\ e \sim \mathrm{N}_n(0, \sigma^2 V) \end{cases}$$

에서, V는 알려져 있는 행렬로서 양의 정부호(positive definite) 행렬이고 주어진 상수의 행렬 X는 $n \times (p+1)$ 행렬로서 계수(rank)가 $(p+1)$이다. 이러한 모형에서는

$$\hat{\beta} = (X^t V^{-1} X)^{-1} X^t V^{-1} \mathrm{Y}$$

로 정의되는 표본회귀계수와 σ^2의 추측값으로는

$$\widehat{\sigma^2} = (\mathrm{Y} - X\hat{\beta})^t V^{-1} (\mathrm{Y} - X\hat{\beta}) / (n-p-1)$$

을 사용하여 추론을 하게 된다. 이 때 다음에 답하여라.

(a) $\hat{\beta}$ 의 분포를 구하여라.

(b) $(n-p-1)\widehat{\sigma^2}/\sigma^2$ 의 분포를 구하여라.

(c) $\hat{\beta}$ 와 $\widehat{\sigma^2}$ 이 서로 독립임을 밝혀라.

[6](20점: (a)10점 (b)10점)

$(\mathrm{X}_1, \mathrm{Y}_1)', \cdots, (\mathrm{X}_n, \mathrm{Y}_n)'$ $(n > 2)$이 이변량 정규분포

$$N_2\left(\begin{pmatrix}\mu_1 \\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right),\ \sigma_1^2 > 0,\ \sigma_2^2 > 0,\ -1 < \rho < 1$$

에서의 랜덤표본이고, r이 표본상관계수 즉

$$r = \frac{\sum_1^n (\mathrm{X}_i - \overline{\mathrm{X}})(\mathrm{Y}_i - \overline{\mathrm{Y}})}{\sqrt{\sum_1^n (\mathrm{X}_i - \overline{\mathrm{X}})^2}\sqrt{\sum_1^n (\mathrm{Y}_i - \overline{\mathrm{Y}})^2}}$$

이라고 할 때, 다음에 답하여라.

(a) 표본상관계수의 표본분포에 대하여 다음의 대의적 정의가 성립하는 것을 밝혀라.

$$\frac{r}{\sqrt{1-r^2}} \overset{d}{\equiv} \frac{\sqrt{\mathrm{V}_1}\,\rho/\sqrt{1-\rho^2} + \mathrm{Z}}{\sqrt{\mathrm{V}_2}}$$

여기에서 $\mathrm{V}_1 \sim \chi^2(\mathrm{n}-1), \mathrm{V}_2 \sim \chi^2(\mathrm{n}-2), \mathrm{Z} \sim \mathrm{N}(0,1)$ 이고 $\mathrm{V}_1, \mathrm{V}_2, \mathrm{Z}$는 서로 독립이다.

(b) 다음의 확률이 $|\rho|$의 증가함수임을 밝히고, 이 확률의 최소값을 구하여라.

$$\mathrm{P}_\rho\left\{\sqrt{n-2}\,\frac{|r|}{\sqrt{1-r^2}} \geq t_{0.025}(n-2)\right\}$$