**Daeyoung Lim**
**Grad School**
**2015, Late**

1. Let $X$ and $Y$ be random variables for which the joint PDF is as follows:

$$f(x, y) = \begin{cases} xe^{-y}, & \text{for } 0 < x < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

   (a) Compute $\mathrm{E}(Y \mid X)$

   (b) Compute $\mathrm{E}\left(e^{tY} \mid X\right)$

   (c) Let $W = Y - X$. Find the PDF of $W$.

**Solution:**

(a) 먼저 $X$의 주변분포는 적분하면 $f_X(x) = xe^{-x}$. 따라서 조건부 분포는

$$f_{Y|X}(y|x) = e^{x-y} \tag{1}$$

이며 정의에 따라

$$\begin{aligned} \mathrm{E}(Y \mid X = x) &= \int_x^{\infty} ye^{x-y}\, dy && (2) \\ &= e^x \int_x^{\infty} ye^{-y}\, dy && (3) \\ &= e^x \left( \left[ -ye^{-y} \right]_x^{\infty} + \int_x^{\infty} e^{-y}\, dy \right) && (4) \\ &= e^x \left( xe^{-x} + e^{-x} \right) && (5) \\ &= x + 1 && (6) \\ \mathrm{E}(Y \mid X) &= X + 1 && (7) \end{aligned}$$

(b) 정의와 *the law of unconscious statistician*에 의해

$$\begin{aligned}
\mathrm{E}\left(e^{tY} \mid X = x\right) &= \int_x^\infty e^{ty} e^{x-y}\, dy && (8)\\
&= e^x \int_x^\infty e^{(t-1)y}\, dy && (9)\\
&= e^x \left[\frac{1}{t-1} e^{(t-1)y}\right]_x^\infty && (10)\\
&= \begin{cases} \dfrac{1}{t-1} e^{tx}, & \text{if } t < 1 \\ \infty, & \text{if } t \geq 1 \end{cases} && (11)\\
\mathrm{E}\left(e^{tY} \mid X\right) &= \begin{cases} \dfrac{1}{t-1} e^{tX}, & \text{if } t < 1 \\ \infty, & \text{if } t \geq 1 \end{cases} && (12)
\end{aligned}$$

(c) $W = Y - X$, $V = X$로 두면 $W \sim \mathrm{Exp}\,(1)$이다.

2. $X_1, \ldots, X_n$의 평균이 $\theta$인 지수분포에서 뽑은 임의표본이라고 하자.

(a) 이 지수분포의 분산에 대한 최대가능도 추정량(MLE)의 점근적 분산 (asymptotic variance)을 구하시오.

(b) $n = 30$이고 표본평균이 26.5라고 할 때, $\Pr\,(X > 10)$의 최대가능도 추정량(MLE)을 이용하여 95% 근사 신뢰구간을 구하시오.

(c) 이 분포로부터 20에서 중도 절단된(right-censored) 크기 5인 표본 $7, 12, 15, 20, 20$이 주어졌다고 할 때, $\theta$의 최대가능도 추정치를 구하시오.

**Solution:**

(a) MLE의 점근분포는 다음과 같다.

$$\widehat{\theta} \xrightarrow{d} \mathcal{N}\left(\theta, \mathcal{I}\,(\theta)^{-1}\right) \tag{13}$$

따라서 정보량을 구하면

$$\mathrm{I}\,(\theta) = \theta^{-2} \tag{14}$$

따라서

$$\widehat{\theta} \xrightarrow{d} \mathcal{N}\left(\theta, n\theta^{-2}\right) \tag{15}$$

(b) 먼저 확률을 구하자.

$$\Pr(X > 10) = \int_{10}^{\infty} \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right) \tag{16}$$
$$= \exp\left(-\frac{10}{\theta}\right) \tag{17}$$
$$\Pr\widehat{(X > 10)} = \exp\left(-\frac{10}{\widehat{\theta}}\right), \qquad \text{(By the invariance property)} \tag{18}$$

*Delta method*를 사용하면 다음과 같다.

$$\sqrt{n}\left(g\left(\widehat{\theta}\right) - g(\theta)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 \left[g'(\theta)\right]^2\right) \tag{19}$$

따라서 $g\left(\widehat{\theta}\right) = \Pr\widehat{(X > 10)}$이므로

$$\sqrt{n}\left(g\left(\widehat{\theta}\right) - g(\theta)\right) \xrightarrow{d} \mathcal{N}\left(0, n\theta^{-2} \cdot \frac{100}{\theta^4} \exp\left(-\frac{20}{\theta}\right)\right) \tag{20}$$

이제 편의를 위해 $\sigma = 10\sqrt{n}\theta^{-3} \exp\left(-10\theta^{-1}\right)$라 하자. 그러면

$$\frac{\sqrt{n}\left(g\left(\widehat{\theta}\right) - g(\theta)\right)}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1) \tag{21}$$

그런데 신뢰구간을 구할 때 $\sigma$에 있는 모수 $\theta$가 매우 계산을 복잡하게 한다. 이는 MLE인 $\widehat{\theta}$로 대체해도 된다. 그렇게 대체한 것을 $\widehat{\sigma}$라 하자. 왜냐하면 *Slutsky's theorem*에 의해

$$\frac{\sqrt{n}\left(g\left(\widehat{\theta}\right) - g(\theta)\right)}{\widehat{\sigma}} \xrightarrow{d} \frac{\sqrt{n}\left(g\left(\widehat{\theta}\right) - g(\theta)\right)}{\sigma} \tag{22}$$

이기 때문이다. 그러면 신뢰구간은 다음과 같이 나온다.

$$g\left(\widehat{\theta}\right) - 1.96\frac{\widehat{\sigma}}{\sqrt{n}} \leq g(\theta) \leq g\left(\widehat{\theta}\right) + 1.96\frac{\widehat{\sigma}}{\sqrt{n}} \tag{23}$$

문제에서 $\widehat{\theta} = 26.5$이고 $n = 30$이라 했으므로 $g(26.5) = e^{-10/26.5}$, $\widehat{\sigma} = 10\sqrt{30} \cdot 26.5^{-3} \cdot e^{-10/26.5}$이므로

$$e^{-10/26.5}\left(1 - \frac{19.6}{26.5^3}\right) \leq g(\theta) \leq e^{-10/26.5}\left(1 + \frac{19.6}{26.5^3}\right) \tag{24}$$

3. 이산형 확률변수 $X$의 확률밀도함수가 로 주어졌다고 하자.

| $X$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $f(x;\theta_0)$ | 0.02 | 0.03 | 0.05 | 0.30 | 0.50 | 0.10 |
| $f(x;\theta_1)$ | 0.50 | 0.30 | 0.10 | 0.05 | 0.03 | 0.02 |

$$H_0 : \theta = \theta_0 \quad \text{대} \quad H_1 : \theta = \theta_1$$

을 고려할 때

(a) 유의수준이 0.10인 모든 기각역을 제시하라.

(b) 위의 확률밀도함수로부터 얻은 하나의 랜덤표본인 $x$의 값이 6이라 하자. 문제 (a)에서 구한 기각역의 검정력을 비교하여 유의수준이 0.10인 최강력 검정법을 시행하라.

**Solution:**

(a) 기각역이 2개 있다.

$$\mathrm{RR}_1 = \{1, 2, 3\}, \quad \mathrm{RR}_2 = \{6\} \tag{25}$$

(b) 두 기각역에 대하여 검정력을 비교하면

$$\Pr(\mathrm{RR}_1 \mid H_1) = 0.9 \tag{26}$$

$$\Pr(\mathrm{RR}_2 \mid H_1) = 0.02 \tag{27}$$

따라서 유의수준 0.10인 검정에서 최강력 검정법의 기각역은 $\{1, 2, 3\}$이므로 관측값이 6인 경우 귀무가설은 기각되지 않는다.

4. 선형회귀모형(Linear Regression Model) $Y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \epsilon_i$을 고려하자. 여기서 $\epsilon_i$는 회귀모형 기본가정을 만족하고 $i = 1, 2, \ldots, n$이다.

(a) 회귀모형 기본가정을 모두 나열하고, 어떻게 기본가정 만족 여부를 확인하는지 간단히 서술하시오.

(b) 위 모형에서 회귀계수모수 $\beta_0, \beta_1, \ldots, \beta_k$와 분산모수 $\sigma^2$의 최대우도추정(Maximum Likelihood Estimation; MLE) 방법에 의해 구하시오.

(c) 위 모형에서 회귀계수모수 $\beta_0, \beta_1, \ldots, \beta_k$와 분산모수 $\sigma^2$를 최소제곱추정(Least Squares Estimation; LSE) 방법에 의해 구하시오.

(d) 위 모형에서 분산이 일정하지 않은 경우, 즉, $\mathrm{Var}(Y_i) = \sigma_i^2$, 회귀계수모수 $\beta_0, \beta_1, \ldots, \beta_k$를 가중최소제곱추정(Weighted Least Squares Estimation; WLSE) 방법에 의해 구하시오.

**Solution:**

(a) 회귀모형의 기본 가정은 다음과 같다.

- linearity and additivity: 반응변수 $Y$의 기댓값은 설명변수 $X_i$들의 선형결합으로 표현될 수 있다.
- statistical independence: 오차항들은 통계적으로 독립이다. 특히 자기상관(autocorrelation)이 없어야 한다.
- homoskedasticity: 오차항들은 모두 같은 분산을 가진다.
- normality: 오차항들은 모두 정규분포를 따른다.

각각의 가정에 대해 검정하는 방법은 차례대로

- linearity and additivity: 산점도행렬을 그려본다. 또한 적합값(fitted values)들과 반응변수의 산점도를 그려본다.
- statistical independence: 통계적으로 독립인지는 모형을 세울 때의 가설이므로 검정하기는 어려울 수 있다. 자기상관성은 *Durbin-Watson test*로 검정할 수 있다.
- homoskedasticity: 산점도 행렬을 통해 검정할 수 있다.
- normality: Q-plot을 그려본다.

이밖에도 다중공선성(multicollinearity)가 거의 없어야 한다. 이는 설계행렬(design matrix) $\mathbf{X}$의 열들이 선형독립이어야 한다는 의미이다. 그렇지 않으면 회귀계수의 추정치가 유일하지 않다. 다중공선성은 $\mathbf{X}'\mathbf{X}$(Grammian matrix)의 condition number를 통해 확인할 수 있다.

(b) 다변량 정규분포를 이용하자. $Y = \mathbf{X}\beta + \epsilon$, $\epsilon \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}_n\right)$. 따라서

$$L\left(\beta, \sigma^2 \,|\, Y, \mathbf{X}\right) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\left(Y - \mathbf{X}\beta\right)'\left(Y - \mathbf{X}\beta\right)\right) \tag{28}$$

따라서 로그 가능도함수는

$$\ell\left(\beta, \sigma^2 \,|\, Y, \mathbf{X}\right) = -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(Y - \mathbf{X}\beta\right)'\left(Y - \mathbf{X}\beta\right) \tag{29}$$

$$\frac{\partial}{\partial\beta}\ell = -\frac{1}{2\sigma^2}\left(2\mathbf{X}'\mathbf{X}\beta - 2\mathbf{X}'Y\right) = 0 \tag{30}$$

$$\widehat{\beta}^{\mathrm{MLE}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'Y \tag{31}$$

$$\frac{\partial}{\partial\sigma^2}\ell = -\frac{n}{2\sigma^2} + \frac{1}{2\left(\sigma^2\right)^2}\left(Y - \mathbf{X}\beta\right)'\left(Y - \mathbf{X}\beta\right) \tag{32}$$

$$\widehat{\sigma}^2_{\mathrm{MLE}} = \frac{1}{n}\left(Y - \mathbf{X}\widehat{\beta}^{\mathrm{MLE}}\right)'\left(Y - \mathbf{X}\widehat{\beta}^{\mathrm{MLE}}\right) \tag{33}$$

(c) 이건 알아서...

(d) WLSE는 GLS의 특수한 형태이다. GLS에서 오차의 공분산 행렬을 어떤 $\sigma^2\mathbf{V}$라고 두는데 여기서 행렬 $\mathbf{V}$에 특별한 제약이 없다. 그러나 WLSE는 오차의 공분산행렬이 대각행렬이어야 한다. 즉, $\epsilon \sim \mathcal{N}\left(\mathbf{0}, \mathbf{W}\right)$이며 $\mathbf{W} = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\right)$이다. GLS에서도 변환을 통해 일반적인 선형회귀 모형으로 돌려놓는 데에 행렬 $\mathbf{V}$의 *square-root matrix*를 사용했다. WLSE에서도 똑같이 $\mathbf{W}$행렬의 *square-root matrix* $\mathbf{W}^{-1/2} = \mathrm{diag}\left(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1}\right)$를 사용한다. 즉,

$$\mathbf{W}^{-1/2}Y = \mathbf{W}^{-1/2}\mathbf{X}\beta + \mathbf{W}^{-1/2}\epsilon \tag{34}$$

$$\mathbf{W}^{-1/2}\epsilon \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}_n\right) \tag{35}$$

가 되어 다시 최소제곱법을 쓸 수 있게 된다. 즉,

$$\widehat{\beta}^{\mathrm{WLSE}} = \underset{\beta}{\operatorname{argmin}} \left(\mathbf{W}^{-1/2}Y - \mathbf{W}^{-1/2}\mathbf{X}\beta\right)' \left(\mathbf{W}^{-1/2}Y - \mathbf{W}^{-1/2}\mathbf{X}\beta\right) \tag{36}$$

이는 다시

$$\left(Y - \mathbf{X}\beta\right)' \mathbf{W}^{-1} \left(Y - \mathbf{X}\beta\right) = \beta'\mathbf{X}'\mathbf{W}^{-1}\mathbf{X}\beta - 2\beta'\mathbf{X}'\mathbf{W}^{-1}Y + \mathrm{const} \tag{37}$$

$$\widehat{\beta}^{\mathrm{WLSE}} = \left(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{W}^{-1}Y \tag{38}$$

가 된다.