Daeyoung Lim

Grad School

Year of 2012, late

1. If X and Y are independent exponential random variables with respective means $1/\lambda_1$ and $1/\lambda_2$.

   (a) Compute the distribution of $Z = \min(X, Y)$.

   (b) What is the conditional distribution of Z given that $Z = X$?

   (c) Consider the function $\lambda_X(t)$ defined as follows:

   $$\exp\left[-\int_0^x \lambda_X(t)\, dt\right] = 1 - F_X(x), \tag{1}$$

   where $F_X(x) = \Pr(X \le x)$. The function $\lambda_X(t)$ is called the failure rate function of X. Show that the failure rate function of X is constant for $t > 0$.

   (d) Show that

   $$\Pr(X < Y \mid Z = t) = \frac{\lambda_X(t)}{\lambda_X(t) + \lambda_Y(t)}, \tag{2}$$

   where $\lambda_X(t)$ and $\lambda_Y(t)$ are the failure rate functions of X and Y.

Solution: Rate parameter로 표기하면 $X \sim \operatorname{Exp}(\lambda_1)$, $Y \sim \operatorname{Exp}(\lambda_2)$이다.

(a) CDF부터 시작하면

$$\begin{aligned}
\Pr(\min(X,Y) \le z) &= 1 - \Pr(X > z, Y > z) && (3)\\
&= 1 - \Pr(X > z)\Pr(Y > z) && (4)\\
&= 1 - \left(\int_z^\infty \lambda_1 e^{-\lambda_1 x}\, dx \int_z^\infty \lambda_2 e^{-\lambda_2 y}\, dy\right) && (5)\\
&= 1 - e^{-\lambda_1 z} e^{-\lambda_2 z} && (6)\\
&= 1 - e^{-(\lambda_1+\lambda_2)z} && (7)\\
\frac{d}{dz}\Pr(\min(X,Y) \le z) &= (\lambda_1 + \lambda_2)\, e^{-(\lambda_1+\lambda_2)z}, \quad z > 0 && (8)
\end{aligned}$$

그러므로 $Z \sim \operatorname{Exp}(\lambda_1 + \lambda_2)$이다.

(b) 정의대로 계산하면 된다.

$$\begin{aligned}
\Pr(Z \le z \mid Z = X) &= 1 - \Pr(Z \ge z \mid Z = X) && (9)\\
&= 1 - \frac{\Pr(Z \ge z, X < Y)}{\Pr(X < Y)} && (10)\\
&= 1 - \frac{\Pr(Z \le X < Y)}{\Pr(X < Y)} && (11)\\
\Pr(X < Y) &= \int_0^\infty \int_0^y \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 y}\, dx\, dy && (12)\\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2} && (13)\\
\Pr(Z \le X < Y) &= \int_z^\infty \int_z^y \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 y}\, dx\, dy && (14)\\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2} e^{-(\lambda_1+\lambda_2)z} && (15)\\
\therefore \Pr(Z \le z \mid Z = X) &= 1 - e^{-(\lambda_1+\lambda_2)z} && (16)\\
Z \mid Z = X &\sim \mathrm{Exp}(\lambda_1 + \lambda_2) && (17)
\end{aligned}$$

where $\mathrm{Exp}(\lambda_1 + \lambda_2)$ denotes the exponential distribution whose mean is $(\lambda_1 + \lambda_2)^{-1}$.

(c) 위에서 풀어봤으면 알겠지만 $X \sim \mathrm{Exp}(\lambda_1)$인 X의 CDF는 $1 - e^{-\lambda_1 x}$, 고로 $1 - F_X(x) = e^{-\lambda_1 x}$이다. 즉, 미적분의 기본정리에 의해

$$\frac{d}{dx} \int_0^x \lambda_X(t) = \lambda_X(x) = \lambda_1 \tag{18}$$

로 상수가 나오게 된다.

(d) 이것도 그냥 정의대로 풀면 된다.

$$\begin{aligned}
\Pr(X < Y \mid Z = t) &= \frac{\Pr(X < Y, Z = t)}{\Pr(Z = t)} && (19)\\
&= \frac{\Pr(X < Y, X = t)}{\Pr(Z = t)} && (20)\\
&= \frac{\Pr(X = t, t < Y)}{\Pr(Z = t)} && (21)\\
&= \frac{\Pr(X = t)\Pr(t < Y)}{\Pr(Z = t)}, \qquad \text{적분은 스스로...} && (22)\\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2} && (23)\\
&= \frac{\lambda_X(t)}{\lambda_X(t) + \lambda_Y(t)} && (24)
\end{aligned}$$

2. One observation is taken on a discrete random variable Y with a probability mass function $f(y \mid \theta) = \Pr(Y = y \mid \theta)$, where $\theta \in \{-1, 0, 1\}$. Find the maximum likelihood estimate (MLE) of θ.

Y	$\Pr(Y = y \mid \theta = -1)$	$\Pr(Y = y \mid \theta = 0)$	$\Pr(Y = y \mid \theta = 1)$
1	1/3	1/4	0
2	1/3	1/4	0
3	0	1/4	1/4
4	1/6	1/4	1/2
5	1/6	0	1/4

Solution: 표까지 그려가면서 정성스럽게 쉬운 문제를 냈는데, 그냥 각 관측값마다 확률이 (정확하게는 *unnormalized likelihood function*이) 가장 큰 모수값을 선택해주면 된다. 그러니까

$$\widehat{\theta}^{\mathrm{MLE}} = \begin{cases} -1, & \text{if } Y = 1 \text{ or } Y = 2 \\ 0 \text{ or } 1, & \text{if Y=3} \\ 1, & \text{if } Y = 4 \text{ or } Y = 5 \end{cases} \tag{25}$$

3. $X_1, X_2, \ldots, X_{100}$이 포아송 분포 $\mathrm{Poi}(\lambda)$로부터 추출한 랜덤 표본이라 하자. $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1 (> \lambda_0)$에 대한 검정을 고려하자.

(a) 균일최강력검정법(uniformly most powerful test)의 기각 영역을 구하시오.

(b) (a)에서 구한 검정법의 검정력 함수(power function)를 구하시오.

(c) 대립가설이 $H_1 : \lambda > \lambda_0$이라면 균일최강력검정법(uniformly most powerful test)이 존재하는지 보이시오.

Solution:

(a) 가능도비를 구한다.

$$\frac{L_0}{L_1} = \frac{e^{-\lambda_0 n}\lambda_0^{X_1+\cdots+X_n}}{e^{-\lambda_1 n}\lambda_1^{X_1+\cdots+X_n}} \tag{26}$$

$$= e^{(\lambda_1-\lambda_0)n}\left(\frac{\lambda_0}{\lambda_1}\right)^{X_1+\cdots+X_n} \leq k \tag{27}$$

일 때 귀무가설이 기각되는 것이 균일최강력검정법이다. $\lambda_1 > \lambda_0$이므로 λ_0의 λ_1에 대한 비는 1보다 작다. 따라서 기각영역은

$$\sum_{i=1}^{n} X_i \geq c \tag{28}$$

일 때이며 (본 문제에서는 $n = 100$), 상수 c는 다음과 같이 구한다.

$$\Pr\left(\sum_{i=1}^{n} X_i \geq c \,\middle|\, H_0\right) = \alpha \tag{29}$$

여기서 $X_1 + \cdots + X_n \sim \text{Poi}\,(n\lambda)$이므로 (29)는

$$\sum_{x=\lceil c \rceil}^{\infty} \frac{e^{-n\lambda_0}\left(n\lambda_0\right)^x}{x!} = \alpha \tag{30}$$

가 된다. 따라서 저 식을 만족하는 실수 c를 찾아서 그것보다 모든 변수를 다 더한 값이 크면 귀무가설을 기각한다.

(b) 검정력 함수는 대립가설을 참으로 놓고 기각역의 확률을 구하는 것이므로

$$\beta\left(X_1, \ldots, X_n\right) = \sum_{x=\lceil c \rceil}^{\infty} \frac{e^{-n\lambda_1}\left(n\lambda_1\right)^x}{x!} \tag{31}$$

이며 어차피 c는 정수로 나올 것이므로 $\lceil c \rceil = c$가 될 것이다. 그리고 $n = 100$이다.

(c) 결국 위에서 사용한 대립가설의 정보는 $\lambda_1 > \lambda_0$뿐이므로 대립가설이 그에 따라 바뀌어도 결과는 변함이 없다. 고로 균일최강력검정은 존재하고 그 기각역 및 모든 결과는 위와 동일하다.

4. 단순선형회귀모형에 따르는 n개의 자료

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \ldots, n \tag{32}$$

들을 고려하자.

(a) 식 (1)을 행렬식

$$Y = \mathbf{X}\beta + \epsilon \tag{33}$$

으로 표현하고자 한다. 이 경우 $Y, \mathbf{X}, \beta$ 및 ϵ을 행렬 또는 벡터로 표현하시오. (5점)

(b) 모수 β에 대한 최소제곱추정량(least squares estimator) $\widehat{\beta}$을 행렬대수를 이용하여 정의하고, $\widehat{\beta}$

이 정규방정식

$$\mathbf{X}'\mathbf{X}\widehat{\beta} = \mathbf{X}'Y \tag{34}$$

을 만족함을 보이시오. (5점)

(c) 행렬 $\mathbf{X}$가 full rank일 충분조건을 쓰고 이 경우 최소제곱추정량을 구체적으로 구하시오. (5점)

(d) 행렬 $\mathbf{X}$가 full rank라는 가정 하에서 기울기 모수 β_1에 대한 95% 신뢰구간을 제시하되 신뢰수준이 95%가 될 조건을 나열하시오. (5점)

(e) 위에서 제시한 신뢰구간의 신뢰수준이 95%가 됨을 증명하시오. (5점)

Solution:

(a) 예년도 문제 참조.

(b) 이 문제는 두 가지 방법으로 풀 수 있다. 하나는 가장 편하게 벡터미분을 이용하는 것이다. 즉, 문제를 수치최적화 문제로 치환한 뒤 변수벡터에 대해 미분한 후 0으로 놓고 정리하는 식이다.

$$\widehat{\beta} = \underset{\beta}{\operatorname{argmin}} \left(Y - \mathbf{X}\beta\right)' \left(Y - \mathbf{X}\beta\right) \tag{35}$$

$$\frac{d}{d\beta}\left(Y - \mathbf{X}\beta\right)' \left(Y - \mathbf{X}\beta\right) = -2\mathbf{X}'Y + 2\mathbf{X}'\mathbf{X}\beta = 0 \tag{36}$$

$$\left(\mathbf{X}'\mathbf{X}\right)\widehat{\beta} = \mathbf{X}'Y \tag{37}$$

따라서 최소제곱추정량은 이를 반드시 만족함을 알 수 있다. 다만 이 식이 반드시 성립하는지는 증명해야 한다. 이는 다음과 같이 증명한다. 행렬의 열공간(column space)는 $\mathbf{col}\left(\mathbf{X}\right)$로 영공간(null space)는 $\mathbf{null}\left(\mathbf{X}\right)$로 표현하기로 한다.

Lemma 1 *For any $m \times n$ matrix A, $\mathbf{null}\left(A'A\right) = \mathbf{null}\left(A\right)$.*

Proof: 벡터 x가 $\mathbf{null}\left(A\right)$에 있다고 하자. 즉, $Ax = 0$. 여기서 $A'Ax = 0$도 성립함은 쉽게 알 수 있다. 반대로 만약 $x \in \mathbf{null}\left(A'A\right)$라면 $A'Ax = 0$이다. 다시 여기에 x'를 곱하면 $x'A'Ax = 0$는 변함이 없고 이는 다시 $\left(Ax\right)'\left(Ax\right) = 0$이 되어 Ax벡터의 내적이 된다. 내적이 0이기 위해서는 $Ax = 0$이어야 하므로 $x \in \mathbf{null}\left(A\right)$. 따라서 $x \in \mathbf{null}\left(A'A\right) \iff x \in \mathbf{null}\left(A\right)$. QED

Lemma 2 *For any $m \times n$ matrix A, $\mathbf{col}\left(A'\right) = \mathbf{col}\left(A'A\right)$.*

Proof: 임의의 $x \in \mathbb{R}^n$ 벡터를 택하면 $A'Ax$는 $\mathbf{col}\left(A'A\right)$에 속한다. 이는 다시 $A'\left(Ax\right)$이므로 $\mathbf{col}\left(A'\right)$에도 속한다. 따라서 $\mathbf{col}\left(A'A\right) \subset \mathbf{col}\left(A'\right)$이다. 이제 열공간의 차원을 비교하면 되는데 우회적으로 앞서 *lemma*로 증명했던 영공간을 통해 $\mathbf{col}\left(A'A\right) = \mathbf{col}\left(A'\right)$임을

보일 수 있다. 왜냐하면 $\mathbf{null}\,(A'A) = \mathbf{null}\,(A)$이고 $A'A$와 A모두 열의 개수가 n개로 동일하기 때문에 *rank-nullity theorem*에 의해 열공간의 차원 역시 동일함을 알 수 있다. 다시 *rank theorem*에 따라 $\mathbf{col}\,(A')$의 차원과 $\mathbf{col}\,(A)$의 차원이 동일하므로 $\mathbf{col}\,(A'A)$의 차원과 $\mathbf{col}\,(A')$의 차원이 같다. 따라서 $\mathbf{col}\,(A'A) = \mathbf{col}\,(A')$이다.

QED

이제 다시 돌아와서 $\mathbf{X}'\mathbf{X}\widehat{\beta} = \mathbf{X}'Y$가 항상 성립하는가에 대한 질문에 답해보자. 분명 위에 따르면 $\mathbf{X}'Y \in \mathbf{col}\,(\mathbf{X}')$이며 이는 곧 $\mathbf{col}\,(\mathbf{X}'\mathbf{X})$이므로 언제나 성립함을 알 수 있다. 단, $\widehat{\beta}$가 유일하게 결정되기 위해서는 $\mathbf{X}'\mathbf{X}$의 역행렬이 존재해야 하고 (일반적으로 행보다 열이 많은 $\mathbf{X}$를 가정하므로) $\mathbf{rank}\,(\mathbf{X}'\mathbf{X}) = \#$ of independent columns이기 때문에 $\mathbf{X}$의 열이 모두 선형독립이어야 함을 알 수 있다.

(c) (b) 참조.

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

(d) $\widehat{\beta} \sim \mathcal{N}\left(\beta, \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right)$이므로 $\mathbf{C} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}$라 놓으면 $\mathrm{Var}\left(\widehat{\beta}_1\right) = \sigma^2\mathbf{C}_{22}$. 단순선형회귀의 경우

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \tag{39}$$

$$= \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \tag{40}$$

$$\left(\mathbf{X}'\mathbf{X}\right)^{-1} = \frac{1}{n\sum_{i=1}^n x_i^2 - \left(n\overline{x}_n\right)^2} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix} \tag{41}$$

여기서

$$n\sum_{i=1}^n x_i^2 - \left(n\overline{x}_n\right)^2 = n\sum_{i=1}^n \left(x_i - \overline{x}_n\right)^2 \tag{42}$$

따라서 $\mathbf{C}_{22} = 1\Big/\left(\sum_{i=1}^n \left(x_i - \overline{x}_n\right)^2\right)$이므로 ($\sigma^2$는 모르므로 비편향추정량인 S^2로 대입한다.)

$$\mathrm{Var}\left(\widehat{\beta}_1\right) = \frac{\sum_{i=1}^n \left(y_i - \hat{y}_i\right)^2}{(n-2)\sum_{i=1}^n \left(x_i - \overline{x}_n\right)^2} \tag{43}$$

그러므로

$$\frac{\widehat{\beta}_1 - \beta_1}{\sqrt{\mathrm{Var}\left(\widehat{\beta}_1\right)}} \sim t_{n-2} \tag{44}$$

를 이용하여

$$\Pr\left(-t_{n-2}\left(0.025\right) \le \frac{\widehat{\beta}_1 - \beta_1}{\sqrt{\mathrm{Var}\left(\widehat{\beta}_1\right)}} \le t_{n-2}\left(0.025\right)\right) = 0.95 \tag{45}$$

$$\mathrm{CI}: \; \widehat{\beta}_1 - t_{n-2}\left(0.025\right)\sqrt{\mathrm{Var}\left(\widehat{\beta}\right)} \le \beta_1 \le \widehat{\beta}_1 + t_{n-2}\left(0.025\right)\sqrt{\mathrm{Var}\left(\widehat{\beta}\right)} \tag{46}$$

(43)을 (46)에 대입한다. 이게 성립하려면 처음에 세웠던 모형이 맞아야 한다. 즉, 오차항이 정규분포를 따라야 하고 모두 같은 분산을 가져야 한다. 또한 **X**와 오차항은 독립이어야 한다.

(e) 뭘 증명하라는 건지 모르겠다. By construction, 신뢰수준은 95%이다.