Daeyoung Lim
Grad School
Year of 2011, late

1. (a) 양의 확률변수(positive random variable) X의 1차 적률(평균)이 존재한다고 가정할 때, 다음의 부등식이 성립함을 보이시오.

$$\Pr(X \geq 1) \leq \mathrm{E}(X) \tag{1}$$

(b) $Z_1, Z_2, \ldots$이 다음과 같은 확률밀도함수를 갖는 연속형 확률분포를 만족하는 분포에서 나온 임의표본이라고 가정하고

$$f(z) = \begin{cases} \lambda e^{-\lambda z}, & z > 0,\ \lambda > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

N을 다음과 같은 확률질량함수를 갖는 이산형 확률변수라고 가정하고

$$\Pr(N = n) = \beta(1-\beta)^{n-1},\ n = 1, 2, \ldots,\ 0 < \beta < 1 \tag{3}$$

$Z_1, Z_2, \ldots$와 N이 서로 독립이라고 가정하자.
이때, (a)의 결과를 이용하여 $X = \sum_{i=1}^{N} Z_i$가 1보다 크거나 같을 확률의 상한을 구하시오. 또한 X의 분산을 구하시오.

Solution:

(a) 마코프 부등식(Markov's inequality)는 증명하는 방법이 꽤 여러가지다. 가장 대표적인 3가지를 소개한다.

- 일반적으로 $a > 0$이라 가정할 때

$$aI_{(X \geq a)} = \begin{cases} a, & \text{if } X \geq a \\ 0, & \text{if } X < a \end{cases} \tag{4}$$

따라서 $aI_{(X\geq a)} \leq X$가 성립함을 알 수 있다. 기댓값 연산자(expectation operator) E은 본질적으로 적분이므로 부등호의 방향이 바뀌지 않는다. 그러므로

$$a\mathrm{E}\left(I_{(X\geq a)}\right) \leq \mathrm{E}(X) \tag{5}$$

이 성립하고 indicator variable의 평균은 그 집합의 확률이 되므로

$$\Pr(X \geq a) \leq \frac{\mathrm{E}(X)}{a} \tag{6}$$

이 된다. 따라서 주어진 문제의 $a = 1$인 경우도 성립한다.

- 두 번째 증명은 기댓값을 적분으로 표현하여 증명한 것이다.

$$\begin{aligned} \mathrm{E}(X) &= \int_0^\infty x f_X(x)\,dx && (7)\\ &= \int_0^a x f_X(x)\,dx + \int_a^\infty x f_X(x)\,dx && (8)\\ &\geq \int_0^a 0\cdot f_X(x)\,dx + \int_a^\infty a f_X(x)\,dx && (9)\\ &= a\Pr(X \geq a) && (10)\end{aligned}$$

(9)에서 두 번째 적분은 적분하는 x의 범위가 $a < x < \infty$이므로 하한인 $x = a$로 고정하게 되면 당연히 (8)의 두 번째 적분보다 작아지게 된다. 따라서 똑같이 나온다.

- 마지막은 Lebesgue theory를 이용해서 증명하는 방법이 있다. 일반적인 르벡 측도 (Lebesgue measure) μ가 있을 때 함수 $f \geq 0$에 대하여 다음과 같은 simple function 을 생각할 수 있다.

$$s(x) = \begin{cases} a, & \text{if } f(x) \geq a \\ 0, & \text{if } f(x) < a \end{cases} \tag{11}$$

그러면 $0 \leq s(x) \leq f(x)$이 되므로 르벡 적분(Lebesgue integral)의 성질에 따라 전체 집합 Ω에 대해

$$\int_\Omega f\,d\mu \geq \int_\Omega s\,d\mu \tag{12}$$

가 성립하고 simple function의 르벡 적분은

$$\int_\Omega s\,d\mu = a\mu\left(\{\omega \in \Omega \,|\, f(\omega) \geq a\}\right) \tag{13}$$

이 되어

$$\mu\left(\{\omega \in \Omega \,|\, f(\omega) \geq a\}\right) \leq \frac{1}{a}\int_\Omega f\,d\mu \tag{14}$$

가 된다. 원래 확률변수는 표본공간에서 실수 집합 $\mathbb{R}$로의 대응관계(mapping)이므로 위 증명에서 정의된 함수 f가 된다.

(b) 먼저 마코프 부등식을 이용하여 상한을 구하려면 기댓값을 알아야 한다. 이 기댓값을 구하는 것도 크게 두 가지 방법이 있다.

- 먼저 정의를 그대로 이용하는 방법이다.

$$
\begin{aligned}
\mathrm{E}(X) &= \mathrm{E}\left(\sum_{i=1}^{N} Z_i\right) && (15)\\
&= \mathrm{E}(NZ_1) && (16)\\
&= \mathrm{E}(N)\,\mathrm{E}(Z_1) && (17)\\
&= \frac{1}{\beta\lambda} && (18)
\end{aligned}
$$

가 된다. 왜냐하면 $N \sim \mathrm{Geo}(\beta)$이고 $Z_i \sim \mathrm{Exp}(\lambda)$이기 때문이다. 다시 여기서 지수분포는 rate parameter로 표기되었다.

- 두 번째는 조건부로 표현하는 것이다. 다시 말해, 원래는 $X \mid N = \sum_{i=1}^{N} Z_i$이기 때문에 온전히 X만의 기댓값을 구하기 위해서는 이중기댓값의 법칙(the law of double expectation)을 이용할 수 있다. $X \mid N \sim \mathrm{Ga}(N, \lambda)$이다. 고로

$$
\begin{aligned}
\mathrm{E}(\mathrm{E}(X \mid N)) &= \mathrm{E}\left(\frac{N}{\lambda}\right) && (19)\\
&= \frac{1}{\lambda}\mathrm{E}(N) && (20)\\
&= \frac{1}{\beta\lambda} && (21)
\end{aligned}
$$

가 된다.

아무튼 이렇게 해서 마코프 부등식은 다음과 같이 표현된다.

$$
\Pr(X \geq 1) \leq \frac{1}{\beta\lambda} \tag{22}
$$

- 또한 X의 분산을 구하기 위해서는 역시 분산의 분해(variance decomposition)를 이용하는 편이 편리하다.

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}(\mathrm{E}(X \mid N)) + \mathrm{E}(\mathrm{Var}(X \mid N)) && (23)\\
&= \mathrm{Var}\left(\frac{N}{\lambda}\right) + \mathrm{E}\left(\frac{N}{\lambda^2}\right) && (24)\\
&= \frac{1}{\lambda^2}\left(\frac{1-\beta}{\beta^2} + \frac{\beta}{\beta^2}\right) && (25)\\
&= \frac{1}{\beta^2\lambda^2} && (26)
\end{aligned}
$$

이다. 여기서 기하분포의 평균은 모수의 역수이고 분산은 $(1-\beta)/\beta^2$이 됨을 기억하자. 까먹었으면 직접 구해도 된다. 귀찮을 뿐.

- 직접 구해도 된다. 즉 $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$이므로

$$\begin{aligned}
\mathrm{E}\left(X^2\right) &= \mathrm{E}\left(\sum_{i=1}^{N} Z_i^2 + \sum_{i\neq j} Z_i Z_j\right) && (27)\\
&= \mathrm{E}(NZ_1) + \mathrm{E}\left(\frac{N(N+1)}{2} Z_1 Z_2\right) && (28)\\
&= \mathrm{E}(N)\,\mathrm{E}(Z_1) + \frac{1}{2}\left(\mathrm{E}\left(N^2\right) + \mathrm{E}(N)\right)\mathrm{E}(Z_1)\,\mathrm{E}(Z_2) && (29)\\
&= \mathrm{E}(N)\,\mathrm{E}(Z_1) + \frac{1}{2}\left(\mathrm{E}\left(N^2\right) + \mathrm{E}(N)\right)\{\mathrm{E}(Z_1)\}^2 && (30)
\end{aligned}$$

알아서 잘 하면 첫 번째와 똑같이 나온다. 귀찮아서 못하겠다.

2. $X_1, \ldots, X_n$이 다음과 같은 확률질량함수(probability mass function)를 갖는 이산형 확률분포로부터의 임의 표본이라고 할 때

$$\Pr(X = x) = \begin{cases} \theta, & x = -1, 0 < \theta < 1 \\ (1-\theta)^2 \theta^x, & x = 0, 1, 2, \ldots \end{cases} \tag{31}$$

모수 θ에 관하여 다음의 두 가지 추정량을 고려하자.

$$T_1(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^{n} I(X_i = -1), \quad T_2(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i I(X_i \geq 0). \tag{32}$$

(a) $n^{-1}\sum_{i=1}^{n} X_i = T_2 - T_1$이 됨을 보이고, 이를 이용하여 두 추정량이 각각 θ의 비편향 추정량(unbiased estimator)임을 보이시오.

(b) 두 추정량의 MSE(mean square error)를 구하고 비교하시오.

(c) 모두 θ의 최대가능도 추정량은 (maximum likelihood estimator) 다음과 같이 주어짐을 보이고

$$\widehat{\theta}_n = \frac{2\sum_{i=1}^{n} I(X_i = -1) + \sum_{i=1}^{n} X_i}{2n + \sum_{i=1}^{n} X_i} \tag{33}$$

$\widehat{\theta}_n^2$이 θ^2의 일치추정량(consistent estimator)이 됨을 보이시오.

Solution:

(a) X_i가 취할 수 있는 값은 0부터가 아니라 $-1, 0, 1, 2, \ldots$이므로 0보다 큰 X_i들을 모두 더했을

때 T_2가 된다. 왜냐하면 $I(X_i \geq 0)$항은 $X_i = -1$일 때만 0이 되고 그 이외에서는 1이 되기 때문에 0보다 크거나 같은 확률변수만 남겨두는 효과를 낳게 된다. 따라서 모두 더한 값은 $X_i = -1$인 횟수만큼 T_2에서 빼주어야 하는데 그 '횟수'가 바로 T_1이다. 횟수는 양수인데 반해 $X_i = -1$은 음수여야 하므로 T_2에서 T_1을 빼주어야 그냥 다 더해서 나눈 값이 된다. 수식으로 말하자면

$$-T_1 = \frac{1}{n}\sum_{i=1}^{n} X_i I(X_i = -1) \tag{34}$$

이므로

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i I(X_i \geq 0)}_{T_2} + \underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i I(X_i = -1)}_{-T_1} = \frac{1}{n}\sum_{i=1}^{n} X_i \underbrace{\left(I(X_i = -1) + I(X_i \geq 0)\right)}_{\text{서로소인 집합이므로 항상 1}} \tag{35}$$

이 되어 $\overline{X}_n = T_2 - T_1$이 된다. 각각의 기댓값은

$$\begin{aligned}
\mathrm{E}(T_1) &= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(I(X_i = -1)) && (36)\\
&= \frac{1}{n}\sum_{i=1}^{n} \Pr(X_i = -1) && (37)\\
&= \frac{1}{n}\cdot n\theta = \theta && (38)\\
\mathrm{E}(T_2) &= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(X_i I(X_i \geq 0)) && (39)\\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{x=0}^{\infty} x(1-\theta)^2\theta^x, \qquad \text{멱급수 계산은 알아서...} && (40)\\
&= \frac{1}{n}\sum_{i=1}^{n} \theta = \theta && (41)
\end{aligned}$$

(b) 둘 모두 비편향 추정량이므로 *variance-bias decomposition*에 의해 MSE를 비교하는 것은 분산을 비교하는 것과 같아진다. 둘의 분산을 구하면 다음과 같다.

$$\begin{aligned}
\mathrm{Var}(T_1) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} I(X_i = -1)\right) && (42)\\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(I(X_i = -1)) && (43)\\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left(\mathrm{E}\left([I(X_i = -1)]^2\right) - [\mathrm{E}(I(X_i = -1))]^2\right) && (44)
\end{aligned}$$

이산형 확률변수의 함수의 기댓값은 예전에 고등학교 때 X와 $\Pr(X)$ 표를 그려서 하나

하나 곱해서 더하던 것을 생각해보면 편하다. 함수꼴의 기댓값 역시 *the law of unconscious statistician*에 의해 그냥 원래 확률을 그대로 써도 된다. 고로 $[I(X_i=-1)]^2$은 언제나 1이므로 그 기댓값은 θ가 되고 $I(X_i=-1)$의 기댓값은 그 자체로도 θ이므로

$$\begin{aligned}\mathrm{Var}(T_1) &= \frac{1}{n^2}n\left(\theta-\theta^2\right) &(45)\\ &= \frac{1}{n}\theta(1-\theta) &(46)\end{aligned}$$

이다. T_2의 분산 역시 같은 방식으로 하면

$$\begin{aligned}\mathrm{E}\left(X_iI(X_i\geq 0)\right) &= \sum_{x=0}^{\infty}x(1-\theta)^2\theta^x=\theta &(47)\\ \mathrm{E}\left(\left[X_iI(X_i\geq 0)\right]^2\right) &= \sum_{x=0}^{\infty}x^2(1-\theta)^2\theta^x=\frac{\theta(\theta+1)}{1-\theta} &(48)\\ \mathrm{Var}(T_2) &= \frac{\theta(\theta+1)}{1-\theta}-\theta^2 &(49)\\ &= \frac{\theta\left(\theta^2+1\right)}{1-\theta} &(50)\end{aligned}$$

둘의 분산을 비교해 보았을 때 T_1의 분산이 더 작으므로 T_1이 MSE의 측면에서 더 좋은 추정량이다.

(c) n개의 확률변수 중 $X_i=-1$인 개수를 m개라 하면 $\sum_{i=1}^{n}I(X_i=-1)=m$이 되고 가능도 함수는 다음과 같다.

$$\begin{aligned}L\left(\theta\mid\{X_i\}_{i=1}^n\right) &= \theta^m(1-\theta)^{2(n-m)}\theta^{\sum_{i=1}^n X_iI(X_i\geq 0)} &(51)\\ \ell\left(\theta\mid\{X_i\}_{i=1}^n\right) &= \left(m+\sum_{i=1}^{n}X_iI(X_i\geq 0)\right)\ln\theta+2(n-m)\ln(1-\theta) &(52)\\ \frac{d}{d\theta}\ell\left(\theta\mid\{X_i\}_{i=1}^n\right) &= \frac{m+\displaystyle\sum_{i=1}^{n}X_iI(X_i\geq 0)}{\theta}-\frac{2(n-m)}{1-\theta} &(53)\\ &= \frac{m+\displaystyle\sum_{i=1}^{n}X_iI(X_i\geq 0)-\theta\left(m+\displaystyle\sum_{i=1}^{n}X_iI(X_i\geq 0)+2(n-m)\right)}{\theta(1-\theta)} &(54)\end{aligned}$$

고로 이를 0으로 놓고 θ에 대해 정리한 뒤 m을 원래대로 복원하면

$$\widehat{\theta}_n = \frac{\sum_{i=1}^{n} I\left(X_i = -1\right) + \sum_{i=1}^{n} X_i I\left(X_i \geq 0\right)}{2n + \sum_{i=1}^{n} X_i I\left(X_i \geq 0\right) - \sum_{i=1}^{n} I\left(X_i = -1\right)} \tag{55}$$

$$= \frac{-\sum_{i=1}^{n} (-1)\, I\left(X_i = -1\right) + \sum_{i=1}^{n} X_i I\left(X_i \geq 0\right)}{2n + \sum_{i=1}^{n} X_i I\left(X_i \geq 0\right) + \sum_{i=1}^{n} (-1)\, I\left(X_i = -1\right)} \tag{56}$$

$$= \frac{-2\sum_{i=1}^{n} (-1)\, I\left(X_i = -1\right) + \sum_{i=1}^{n} X_i I\left(X_i \geq 0\right) + \sum_{i=1}^{n} (-1)\, I\left(X_i = -1\right)}{2n + \sum_{i=1}^{n} X_i} \tag{57}$$

$$= \frac{2\sum_{i=1}^{n} I\left(X_i = -1\right) + \sum_{i=1}^{n} X_i}{2n + \sum_{i=1}^{n} X_i} \tag{58}$$

마지막으로 $\widehat{\theta}_n$은 θ의 일치추정량이다. 왜냐하면 대수의 법칙에 의해

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mathrm{E}\left(X_1\right) \tag{59}$$

인데

$$\mathrm{E}\left(X_1\right) = -\theta + \sum_{x=0}^{\infty} x\left(1-\theta\right)^2 \theta^x = -\theta + \theta = 0 \tag{60}$$

이므로 $\overline{X}_n \xrightarrow{p} 0$이며

$$\frac{1}{n}\sum_{i=1}^{n} I\left(X_i = -1\right) \xrightarrow{p} \mathrm{E}\left(I\left(X_1 = -1\right)\right) \quad \Big(= \Pr\left(X_1 = -1\right)\Big) \tag{61}$$

이므로 θ에 수렴하게 된다. 따라서 *Slutsky's theorem*에 의해서

$$\widehat{\theta}_n \xrightarrow{p} \frac{2\theta + 0}{2 + 0} = \theta \tag{62}$$

이므로 MLE의 불변성(invariance property)에 의해 $\widehat{\theta}_n^2$ 역시 θ^2의 일치추정량이 된다.

3. $X_1, \ldots, X_n$이 $\mathcal{N}\left(\mu, \sigma^2\right)$로부터의 랜덤표본이라 할 때, 모평균에 대한 가설 $H : \mu = \mu_0$ vs $K : \mu \neq \mu_0$에 대한 t 검정법을 가능도비를 사용하여 유도하고, 유도된 검정법의 검정력 및 성질에 대해 논하라.

(σ^2는 알려지지 않았음.)

Solution: μ의 MLE는 $\overline{X}_n$임이 자명하므로 다시 구하지 않겠다. 가능도비를 구하면

$$\frac{L(\mu_0)}{L(\widehat{\mu}^{\text{MLE}})} = \frac{\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu_0)^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n \left(X_i-\overline{X}_n\right)^2\right)} \tag{63}$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n (X_i-\mu_0)^2\right) - \sum_{i=1}^n \left(X_i-\overline{X}_n\right)^2\right) \tag{64}$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n \left(X_i-\overline{X}_n\right)^2 + n\left(\overline{X}_n-\mu_0\right)^2 - \sum_{i=1}^n \left(X_i-\overline{X}_n\right)^2\right)\right) \tag{65}$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left(\overline{X}_n-\mu_0\right)^2\right) \le k \tag{66}$$

일 때 귀무가설이 기각된다. 따라서 $\left(\overline{X}_n-\mu_0\right)^2 \ge k_2$일 때, 다시 고치면

$$\left(\frac{\overline{X}_n-\mu_0}{S}\right)^2 \ge \frac{k_2}{S^2} \tag{67}$$

$$\underbrace{\left(\sqrt{n}\frac{\overline{X}_n-\mu_0}{S}\right)^2}_{T^2} \ge n\frac{k_2}{S^2} \tag{68}$$

이므로 T^2가 충분히 클 때 귀무가설은 기각된다. 다시 말해서 T가 충분히 크거나 충분히 작을 때 귀무가설이 기각된다는 의미이다. 그리고 T는 자유도가 $n-1$인 t-분포를 따르므로 $T \le t_{n-1}(1-\alpha)$이거나 $T \ge t_{n-1}(\alpha)$일 때 α 신뢰수준에서 귀무가설은 기각된다.

4. 단순회귀모형(simple regression model) $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$을 고려하자. 여기서 ϵ_i는 기본회귀모형 가정을 만족하고 $i = 1, 2, \ldots, n$이다.

   (a) 위 모형을 행렬 형식으로 자세히 포현하고, 행렬 형태로 회귀 모수 $\beta = (\beta_0, \beta_1)'$의 최소제곱추정치(least square estimator; LSE)를 구하시오.

   (b) LSE $\widehat{\beta} = \left(\widehat{\beta}_0, \widehat{\beta}_1\right)'$의 분포를 구하시오.

   (c) $\widehat{\beta}$와 S^2가 독립임을 보이시오. 여기서 $S^2 = \mathrm{SSE}/\left(n-2\right) = \left(Y - \mathbf{X}\widehat{\beta}\right)'\left(Y - \mathbf{X}\widehat{\beta}\right)$이다.

Solution:

(a) 위 모형은 행렬식으로 표현하면

$$\underbrace{\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}}_{\beta} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}}_{\epsilon} \tag{69}$$

이 되고 LSE는 $\widehat{\beta} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'Y$이다.

(b) $\widehat{\beta}$는 Y의 선형변환이고 $Y \sim \mathcal{N}\left(\mathbf{X}\beta, \sigma^2\mathbf{I}_n\right)$이므로 정규분포의 선형불변성(affine property)에 의해 어떠한 선형변환꼴도 다시 정규분포를 따르게 되어 있다. 따라서

$$\begin{aligned}
\mathrm{E}\left(\widehat{\beta}\right) &= \mathrm{E}\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'Y\right) && (70)\\
&= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\beta && (71)\\
&= \beta && (72)\\
\mathrm{Var}\left(\widehat{\beta}\right) &= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathrm{Var}\left(Y\right)\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} && (73)\\
&= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\cancel{\mathbf{X}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}} && (74)\\
&= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1} && (75)\\
\widehat{\beta} &\sim \mathcal{N}\left(\beta, \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right) && (76)
\end{aligned}$$

(c) 먼저 $\text{Cov}\left(\widehat{\beta}, Y - \mathbf{X}\widehat{\beta}\right)$을 구하면

$$\begin{aligned}
\text{Cov}\left(\widehat{\beta}, Y - \mathbf{X}\widehat{\beta}\right) &= \text{Cov}\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'Y, \left(\mathbf{I}_n - \mathbf{H}\right)Y\right) && (77)\\
&= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\text{Cov}\left(Y\right)\left(\mathbf{I}_n - \mathbf{H}\right)' && (78)\\
&= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\mathbf{I}_n - \mathbf{H}\right) && (79)\\
&= \mathbf{0} && (80)
\end{aligned}$$

정규분포의 경우 공분산이 0이면 서로 독립이므로 $\widehat{\beta}$와 $Y - \mathbf{X}\widehat{\beta}$는 서로 독립이며 그 함수꼴 역시 서로 독립이므로 $\widehat{\beta}$와 S^2 역시 서로 독립이다.