Daeyoung Lim

Grad School

Year of 2010, early

1. 1. 확률변수 X의 적률생성함수 $M(t)$가 존재한다고 가정할 때, 다음을 보이시오.

$$\begin{cases} \Pr(X \geq a) \leq e^{-ta} M(t), & t > 0 \\ \Pr(X \leq a) \leq e^{-ta} M(t) \end{cases}$$

2. $X_1, \ldots, X_n$이 $\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$를 만족하는 분포에서 나온 임의표본일 때, $S_n = \sum_{i=1}^{n} X_i$에 대하여 다음이 성립함을 1을 이용하여 보이시오.

$$\Pr(S_n \geq a) \leq e^{-ta} e^{nt^2/2}, \quad t > 0$$

3. 어떤 도박꾼이 매 배팅에서 같은 확률로 1원을 따거나 잃는 게임을 하려고 한다고 하자. 매 배팅에서의 결과는 서로 독립이라 할 때, 처음 10번의 배팅에서 적어도 8번 이상 이기는 확률의 상한을 2를 이용하여 구하시오. 또, 이 확률의 정확한 값도 계산하시오.

Solution:

1. 이를 *Chernoff's bound*라고 한다. 마코프 부등식(Markov inequality)로부터 간단하게 도출할 수 있다.

$$\Pr(X \geq a) = \begin{cases} \Pr(tX \geq at) = \Pr\left(e^{tX} \geq e^{at}\right), & t > 0 \\ \Pr(tX \leq at) = \Pr\left(e^{tX} \leq e^{at}\right), & t < 0 \end{cases} \tag{1}$$

$$= \begin{cases} \Pr(X \geq a) \leq e^{-ta} \mathrm{E}\left(e^{tX}\right), & t > 0 \\ \Pr(X \leq a) \leq e^{-ta} \mathrm{E}\left(e^{tX}\right), & t < 0 \end{cases} \tag{2}$$

(1)에서 (2)로 넘어가는 것은 마코프 부등식이다.

2. 이 문제는 표준정규분포를 가정하는 것보다 정확한 분포를 알고 그 분포의 적률생성함수를 이용하는 것이 더 *tight*한 bound를 준다는 것을 증명하는 것이다. 왜냐하면 표준정규확률변수 n개를 더한 변수의 적률생성함수가 $e^{nt^2/2}$이기 때문이다.

$$X \mid 1 \quad -1 \tag{3}$$

$$e^{tX} \mid e^{t} \quad e^{-t} \tag{4}$$

$$\Pr(X = x) \left| \; \frac{1}{2} \quad \frac{1}{2} \right. \tag{5}$$

이로서 다음과 같은 사실을 얻을 수 있다.

$$\mathrm{E}\left(e^{tX}\right) = \frac{1}{2}\left(e^{t} + e^{-t}\right) \quad \Longrightarrow \quad \mathrm{E}\left(e^{tS_n}\right) = \left(\frac{e^{t} + e^{-t}}{2}\right)^{n} \tag{6}$$

따라서 $2^{-n}\left(e^{t} + e^{-t}\right)^{n} \leq e^{nt^2/2}$임을 보이면 된다.

$$e^{t} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \tag{7}$$

$$e^{-t} = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} \tag{8}$$

$$e^{t^2/2} = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{t^2}{2}\right)^{k} = \sum_{k=0}^{\infty} \frac{t^{2k}}{k! \times 2^k} \tag{9}$$

(7)과 (8)을 더하면 k가 홀수인 부분은 소거되고 짝수인 부분만 두 배가 된다. 따라서

$$\frac{1}{2}\left(e^{t} + e^{-t}\right) = \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!} \tag{10}$$

분모를 비교하면 $k! \times 2^k < (2k)!$이므로 증명끝.

3. 8번 이기면 2번은 진 것이므로 $S_n = 6$이고 9번 이기면 1번은 진 것이므로 $S_n = 8$, 10번 모두 이기면 $S_n = 10$이다. 2번을 이용하면

$$\Pr\left(S_n \geq 6\right) \leq e^{5t^2 - 6t} \tag{11}$$

이므로 최대값은

$$5\left(t^2 - \frac{6}{5}t + \frac{9}{25}\right) - \frac{9}{5} = 5\left(t - \frac{3}{5}\right)^2 - \frac{9}{5} \tag{12}$$

$t = 3/5$에서 $e^{-9/5}$가 된다. 정확한 확률값은

$$\Pr\left(S_n = 6\right) + \Pr\left(S_n = 8\right) + \Pr\left(S_n = 10\right) = \left(\binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right)\left(\frac{1}{2}\right)^{10} \tag{13}$$

$$= \frac{7}{128} \tag{14}$$

2. Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform distribution on $\left(\mu - \sqrt{3}\sigma, \mu + \sqrt{3}\sigma\right)$. Here the unknown parameters are μ and σ, which are the population mean and standard deviation.

    1. Obtain the probability density function of X_i.
    2. Obtain the likelihood function of μ and σ.

3. Obtain the maximum-likelihood estimators (MLEs) of μ and σ.
4. Obtain the method-of-moments estimators of μ and σ.

Solution:

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with density

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!},$$

where $\lambda \geq 0$ and $x = 0, 1, 2, \ldots$. Let $T = \sum_{i=1}^{n} X_i$.

1. Show that T is complete and sufficient for λ.
2. Show that T/n is the uniformly minimum variance estimator (UMVUE) of λ.
3. Show that $T(T-1)\cdots(T-k+1)/n^k$ is the UMVUE of λ^k for any positive k.

Solution:

4. 세 변수에 대한 n개의 관찰값 (x_{i1}, x_{i2}, Y_i), $(i = 1, \ldots, n)$에 대하여

$$\mathrm{E}(Y_i \mid x_{i1}, x_{i2}) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}, \text{ 그리고 } \mathrm{Var}(Y_i \mid x_{i1}, x_{i2}) = \sigma^2$$

이라 하자.

1. 모수 $\beta_0, \beta_1, \beta_2$의 최소제곱추정량 $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2$을 유도하시오.
2. $s^2 = \frac{1}{n-3}\sum_{i=1}^{n}\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{i1} - \widehat{\beta}_2 x_{i2}\right)^2$이 σ^2에 대한 불편추정량이 됨을 보이시오.

Solution:

5. X_n이 자유도가 n인 카이제곱분포를 따른다고 한다. 이 때

$$Z_n = \frac{X_n - n}{\sqrt{2n}}$$

이라고 하자. n이 무한대로 접근함에 따라 Z_n의 분포가 어느 분포에 수렴하는지를 자세히 보아라. (힌트: 자유도가 n인 카이제곱확률변수는 자유도가 1인 서로 독립인 카이제곱확률변수의 합으로 표현될 수 있다.)