연습문제 5.12.

- $X_1 \dots X_n \sim Poi(\lambda)$.
- $\sqrt{n}\big(g(\bar{X}) - g(\lambda)\big) \xrightarrow{d} Z, \quad Z \sim N(0,1)$ 을 만족하는 변환 g를 구하라.

- CLT를 쓰면
$$\sqrt{n}(\bar{X}_n - \lambda) \xrightarrow{d} N(0, \lambda)$$
- $g(\cdot)$이 미분가능한 함수라고 하자.
- $g(\cdot)$는 미분가능한 어떠한 형태도 가능하다. 예를들어 아래와 같은 형태들이 가능하다고 하자.
$$g(x) = x^2$$
혹은
$$g(x) = 2\sqrt{x}$$
- 예를들어서 $g(x) = 2\sqrt{x}$와 같은 형태라고 하자. 그렇다면 $x = a$에서 $g(x)$를 테일러 전개하면 아래와 같이 표현가능하다.
$$g(x) = g(a) + (x-a)g'(a) + (x-a)^2 O(1)$$
- x대신에 $\bar{X}_n$ a대신에 λ를 대입하면
$$g(\bar{X}_n) = g(\lambda) + (\bar{X}_n - \lambda)g'(\lambda) + (\bar{X}_n - \lambda)^2 O(1).$$
양변에 $\sqrt{n}$을 곱하고 정리하면
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) = \sqrt{n}(\bar{X}_n - \lambda)g'(\lambda) + \sqrt{n}(\bar{X}_n - \lambda)^2 O(1).$$
그런데 $\sqrt{n}(\bar{X}_n - \lambda)^2 O(1) = O_p(1)o_p(1)O(1) = o_p(1)$ 이므로
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) = \sqrt{n}(\bar{X}_n - \lambda)g'(\lambda) + o_p(1)$$
그런데 $g(x) = 2\sqrt{x}$이므로 $g'(x) = 1/\sqrt{x}$이다. 따라서 $g'(\lambda) = 1/\sqrt{\lambda}$.
- 정리하면
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) = \sqrt{n}(\bar{X} - \lambda)1/\sqrt{\lambda} + o_p(1)$$
CLT의 결과를 쓰면
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) \xrightarrow{d} N(0, \lambda)1/\sqrt{\lambda}$$
따라서
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) \xrightarrow{d} N(0, 1)$$

(결론) 만약에 $g(x) = 2\sqrt{x}$로 설정한다면
$$\sqrt{n}\big(g(\bar{X}) - g(\lambda)\big)$$
의 극한이 표준정규분포로 수렴하므로 문제에서 원하는 답을 구할 수 있다.

- 이는 바꾸어 말하면 포아송분포에 한정하여 (1) 관측한 확률변수에서 제곱근을 취한뒤 2배를 하고 (2) 그상태에서 CLT를 쓴다면 표준정규분포로 수렴한다는 것이다.
- 이 경우 그냥 포아송분포자체에서 CLT를 쓰는것보다 이득되는 경우가 있다. (일단 이렇게만 알아둘것)

- 그렇다면 어떻게 $g(x) = 2\sqrt{x}$임을 알 수 있을까?
- g가 미분가능하다면
$$g(x) \approx g(a) + (x-a)g'(a)$$
이고 다시 정리하면
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) \approx \sqrt{n}(\bar{X}_n - \lambda)g'(\lambda)$$
따라서
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) \xrightarrow{d} N(0, \lambda)g'(\lambda)$$
이고 정규분포의 특성에 따라서
$$\sqrt{n}\big(g(\bar{X}_n) - g(\lambda)\big) \xrightarrow{d} N\big(0, \lambda(g'(\lambda))^2\big)$$
이다. 따라서
$$\lambda(g'(\lambda))^2 = 1$$
이어야 할 조건은
$$g'(\lambda) = \frac{1}{\sqrt{\lambda}}$$
일 조건이다. 즉 $g(x)$의 형태는 모르겠지만 $g'(x)$는 아래와 같아야 함을 알 수 있다.
$$g'(x) = \frac{1}{\sqrt{x}}$$
이러부터 $g(x)$를 유추하면
$$g(x) = \int \frac{1}{\sqrt{x}} dx = 2\sqrt{x}.$$

- 이제 λ의 점근적인 신뢰구간을 구해보자. 아래의 식을 관찰하자.
$$2\sqrt{n}\big(\sqrt{\bar{X}_n} - \sqrt{\lambda}\big) \xrightarrow{d} N(0, 1)$$
이때 $2\sqrt{n}\big(\sqrt{\bar{X}_n} - \sqrt{\lambda}\big)$자체를 하나의 통계량 Y_n으로 본다면
$$Y_n \xrightarrow{d} N(0, 1)$$
Y_n이 정규분포를 따른다면 아래가 성립한다.
$$P\big(-Z_{\alpha/2} < Y_n < Z_{\alpha/2}\big) = 1 - \alpha$$
Y_n이 정규분포는 아니지만 점근적으로 정규분포를 따른다면 아래가 성립한다.
$$\lim_{n\to\infty} P\big(-Z_{\alpha/2} < Y_n < Z_{\alpha/2}\big) = 1 - \alpha$$
따라서
$$\lim_{n\to\infty} P\left(-Z_{\alpha/2} < 2\sqrt{n}(\sqrt{\bar{X}_n} - \sqrt{\lambda}) < Z_{\alpha/2}\right) = 1 - \alpha$$

이는
$$\lim_{n\to\infty} P\left(-Z_{\alpha/2} < 2\sqrt{n}(\sqrt{\lambda}-\sqrt{\bar{X}_n}) < Z_{\alpha/2}\right) = 1-\alpha$$
정리하면
$$\lim_{n\to\infty} P\left(\frac{-Z_{\alpha/2}}{2\sqrt{n}} + \sqrt{\bar{X}_n} < \sqrt{\lambda} < \frac{Z_{\alpha/2}}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right) = 1-\alpha$$
따라서
$$\lim_{n\to\infty} P\left(\left(\frac{-Z_{\alpha/2}}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right)^2 < \lambda < \left(\frac{Z_{\alpha/2}}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right)^2\right) = 1-\alpha$$
이 성립한다.

- 위의 식이 의미하는것은 λ가 구간
$$\left(\left(\frac{-Z_{\alpha/2}}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right)^2, \left(\frac{Z_{\alpha/2}}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right)^2\right)$$
에 포함될 확률이 점근적으로 $1-\alpha$에 수렴한다는 뜻이다. 이는 바꾸어 말하면 매우 큰 n에 대하여 $1-\alpha$의 확률로 모수가 위의 구간에 있음을 확신할 수 있다는 의미이다. 이를 $1-\alpha$의 점근적 신뢰구간이라고 부른다.

- 예를들어 λ의 점근적인 95퍼센트 신뢰구간은 아래와 같이 구할 수 있다.
$$\left(\left(\frac{-1.96}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right)^2, \left(\frac{1.96}{2\sqrt{n}} + \sqrt{\bar{X}_n}\right)^2\right)$$

연습문제 5.14.

- $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \overset{idd}{\sim} F$ with

(1) $EX_1 = EY_1 = 0$,

(2) $EX_1^2 = EY_1^2 = 1$,

(3) $EX_1Y_1 = \rho$,

(4) $EX_1^4 < \infty$, $EY_1^4 < \infty$.

- $\hat{\rho}_n = \frac{\sum_{i=1}^n (X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^n (X_i-\bar{X})^2}\sqrt{\sum_{i=1}^n (Y_i-\bar{Y})^2}}$.

- 아래와 같은 기호를 약속하자.

(1) $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

(2) $\bar{X}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2$.

(3) $\overline{XY} = \frac{1}{n}\sum_{i=1}^n X_iY_i$.

(a) $\sqrt{n}\bar{X}\bar{Y} \overset{p}{\to} 0$ 임을 보여라.

(sol)

- $\sqrt{n}\bar{X} = O_p(1)$, $\bar{Y} = o_p(1)$ 이므로 자명함. 여기에서 $\sqrt{n}\bar{X} = O_p(1)$ 는 CLT에 의해서 성립하고 $\bar{Y} = o_p(1)$ 는 큰수의 법칙 즉 WLLN에 의해서 성립한다.

note: 추가적으로 $\bar{X}Y - \bar{X}\bar{Y} = \bar{X}Y + o_p(\frac{1}{\sqrt{n}})$ 임도 알 수 있다.

(b) 아래가 성립함을 보여라.
$$\sqrt{m_2 - m_1^2} - \sqrt{m_2} = \frac{-m_1^2}{\sqrt{m_2 - m_1^2} + \sqrt{m_2}}$$
이때 $m_1 = \bar{X}$, $m_2 = \bar{X}^2$ 라고 하자.

(sol) (자명하다)

note: 추가적으로 아래를 확인가능하다.
$$\sqrt{m_2 - m_1^2} = \sqrt{m_2} + \frac{-m_1^2}{\sqrt{m_2 - m_1^2} + \sqrt{m_2}} = \sqrt{m_2} + o_p\left(\frac{1}{\sqrt{n}}\right)$$

- $\sqrt{n}m_1^2 = \sqrt{n}m_1m_1 = O_p(1)o_p(1) = o_p(1) \quad \therefore m_1^2 = o_p(\frac{1}{\sqrt{n}})$.

- $m_2 = o_p(1)$ 이고 $m_1^2 = o_p(\frac{1}{\sqrt{n}})$ 이다. $o_p(1) + o_p(\frac{1}{\sqrt{n}}) = o_p(1)$ 임을 이용하면
$$\sqrt{m_2 - m_1^2} + \sqrt{m_2} = o_p(1)$$
이다.

- $\frac{o_p(\frac{1}{\sqrt{n}})}{o_p(1)} = o_p(\frac{1}{\sqrt{n}})$ 가 성립하므로
$$\frac{-m_1^2}{\sqrt{m_2 - m_1^2} + \sqrt{m_2}} = o_p\left(\frac{1}{\sqrt{n}}\right)$$
이다.

(c) 아래가 성립함을 보여라.
$$\sqrt{m_2} - 1 = \frac{1}{2}(m_2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

(sol)

- 아래식을 관찰하라.
$$\sqrt{m_2} - 1 = \frac{m_2 - 1}{\sqrt{m_2} + 1} = \frac{m_2 - 1}{2} + \left(\frac{1}{\sqrt{m_2} + 1} - \frac{1}{2}\right)(m_2 - 1)$$

- 이제 아래를 보이면 된다.
$$\left(\frac{1}{\sqrt{m_2} + 1} - \frac{1}{2}\right)(m_2 - 1) := r_{3n} = o_p\left(\frac{1}{\sqrt{n}}\right)$$
따라서 아래를 보이면 된다.
$$\sqrt{n}\left(\frac{1}{\sqrt{m_2} + 1} - \frac{1}{2}\right)(m_2 - 1) = o_p(1)$$

- 그런데 CLT에 의해서 $\sqrt{n}(m_2 - 1) = O_p(1)$ 가 성립한다.
- continuous mapping thm에 의해서 아래가 성립한다.

$$m_2 \xrightarrow{p} 1 \quad \Longrightarrow \quad \left(\frac{1}{\sqrt{m_2}+1} - \frac{1}{2}\right) \xrightarrow{p} 0$$

- 따라서

$$\left(\frac{1}{\sqrt{m_2}+1} - \frac{1}{2}\right) = o_p(1)$$

- 따라서 증명이 끝난다.

(d) 아래를 보여라.

$$c_1 - m_1 s_1 - \rho\sqrt{m_2 - m_1^2}\sqrt{s_2 - s_1^2}$$

(sol)