제10주 연속확률분포모형(3)

hylee@silla.ac.kr

확률및통계(2)

제 5 절 베타분포

[정의] 베타분포

확률변수 X가 모수$(\alpha, \beta), \alpha > 0, \beta > 0$

를 가지며 다음 확률밀도함수

$$f(x) = \begin{cases} \dfrac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} & (0 < x < 1) \\ 0 & (x \le 0 \text{ 또는 } x \ge 1) \end{cases}$$

를 가질 때, X는베타분포(beta distribution)

를 따른다고 하며 $\beta(\alpha, \beta)$로 표기한다.

여기서, $B(\alpha,\beta)$ 는

$$B(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

와 같은 베타함수(beta function)이다.

***

베타함수와 감마함수의 관계

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

예제 12

베타분포가 확률밀도함수임을 보이시오

[풀이]

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \geq 0 \quad (0 < x < 1)$$

이고,

$$\int_0^1 \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} dx$$

$$= \frac{1}{B(\alpha, \beta)} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$$

$$= \frac{B(\alpha, \beta)}{B(\alpha, \beta)} = 1$$

예제 13

X가 $\alpha = \beta = 2$ 인 베타분포에 따를 때,

$P(X \le 0.1)$를 구하시오.

[풀이]

$\alpha=\beta=2$ 일 때, 베타분포의 확률밀도 함수는 다음과 같다.

$$f(x)=\frac{1}{B(2,2)}x^{2-1}(1-x)^{2-1}$$

$$=\frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}x(1-x)$$

$$= \frac{3!}{1!\,1!} x(1-x) = 6x(1-x), \quad (0 < x < 1)$$

따라서, 구하는 확률은 다음과 같다.

$$P(X \le 0.1) = \int_0^{0.1} 6x(1-x)dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^{0.1}$$

$$= 6\left(\frac{0.01}{2} - \frac{0.001}{3}\right) = 0.028$$

제 6절 χ^2 분포

[정의] χ^2 분포

감마분포에서 $\alpha = n/2(n:$자연수$)$,

$\beta = 1/2$ 일 때

$$f(x) = \frac{1}{2^{1/2}\Gamma(n/2)} x^{n/2} e^{-x/2} \quad (x > 0)$$

로 정의되는 함수를 자유도 n 의 χ^2 분포

(chi - square distributi on) 의 밀도함수

라 한다.

표준정규분포와 χ^2 분포의 관계

[정리]

X 가 $N(\mu,\sigma^2)$ 를 따르면, $Y = \dfrac{(X-\mu)^2}{\sigma^2}$

은 자유도 $d.f.=1$ 인 χ^2 분포를 한다.

[증명]

$Z = \dfrac{X-\mu}{\sigma}$가 $N(0,1)$ 을 따르는 확률변수

이므로, $Y = Z^2$ 이라면, 임의의 양수 y에

대해서 $\sqrt{Y} = Z$ 이므로,

$$P(0 \le Y \le y) = P(-\sqrt{y} \le Z \le \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$z^2 = y$ 로 두면

$$P(0 \le Y \le y) = \frac{2}{\sqrt{2\pi}} \int_0^z \frac{1}{2\sqrt{y}} dy$$

$$= \int_0^z \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} y^{\frac{1}{2}-1} e^{-\frac{y}{2}} dy$$

이다. 따라서 Y 는 자유도가1인 χ^2 분포를 따른다.

예제 14

X가 $N(\mu,\sigma^2)$를 따를때,

$$V^2 = \frac{(\bar{X}-\mu)^2}{\sigma^2/n}$$ 은 자유도 1 인

χ^2 분포를 따름을 보이시오.

[풀이]

$V = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$ 가 $N(0,1)$ 을 따르므로, V^2 은

자유도가1인 χ^2 분포를 따른다.

제 7절 T 분포

[정의] t 분포

X 의 확률밀도함수가

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-\frac{(n+1)}{2}}$$

일 때, 자유도 n인 t분포라 한다.

예제 16

t 분포가 확률밀도함수임을 보이시오.

[풀이]

$$f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma(2/n)}(1+\frac{x^2}{n})^{-\frac{(n+1)}{2}}\geq 0$$

이고,

$$\int_{-\infty}^{\infty}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma(2/n)}(1+\frac{x^2}{n})^{-\frac{(n+1)}{2}}dx$$

$$=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma(2/n)}\int_{-\infty}^{\infty}(1+\frac{x^2}{n})^{-\frac{(n+1)}{2}}dx$$

이므로,

$$\int_{-\infty}^{\infty}(1+\frac{x^2}{n})^{-\frac{(n+1)}{2}}dx = 2\int_{0}^{\infty}(1+\frac{x^2}{n})^{-\frac{(n+1)}{2}}dx$$

이고, $(1+\frac{x^2}{n})^{-1} = t$ 로 치환하면,

$$dx = -\left(\frac{\sqrt{n}}{2}\right)(1-t)^{-\frac{1}{2}}t^{-\frac{3}{2}}dt$$

이다. 따라서,

$$2\int_0^\infty \left(1+\frac{x^2}{n}\right)^{-\frac{(n+1)}{2}} dx = 2\int_1^0 \left(-\frac{\sqrt{n}}{2}\right) t^{\frac{n-2}{2}} (1-t)^{-\frac{1}{2}} dt$$

$$= \sqrt{n}\int_0^1 t^{\frac{n}{2}-1} (1-t)^{\frac{1}{2}-1} dt$$

$$= \sqrt{n}\,B\left(\frac{n}{2},\frac{1}{2}\right) = \sqrt{n}\,\frac{\Gamma({}^{n}\!/_{2})\Gamma({}^{1}\!/_{2})}{\Gamma({}^{n+1}\!/_{2})}$$

$$= \sqrt{n\pi}\,\frac{\Gamma({}^{n}\!/_{2})}{\Gamma({}^{n+1}\!/_{2})}$$

이다. 그러므로,

$$\int_{-\infty}^{\infty} f(x)dx = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n+1}{2})} \cdot \frac{\sqrt{n\pi}\,\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n+1}{2})}$$

$$=1$$

임을 알수있다.

[정리]

확률변수 U, V가 독립이고, U는 표준정규분포 $N(0,1)$를 따르고, V는 자유도 n의 χ^2분포를 따를 때, 확률변수

$$T = \frac{U}{\sqrt{V/n}}$$

는 자유도 n인 t분포를 따른다.

제 8절 F 분포

[정의] F 분포

X 의 확률밀도함수가

$$f(x)=\frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}(1+\frac{n_1}{n_2}x)^{-\frac{1}{2}(n_1+n_2)}$$

$(x>0)$

인 분포를 자유도 (n_1, n_2)인 F 분포라 한다.

[정리]

확률변수 U, V 가 서로 독립이고, 각각 자유도가 n_1 과 n_2 인 χ^2 분포를 따를 때, 확률변수

$$F = \frac{U/n_1}{V/n_2}$$

는 자유도 (n_1, n_2) 인 F 분포를 따른다.

[정리]

확률변수 T가 자유도 n인 t분포를 따를 때, T^2 은자유도가 $(1, n)$인 F 분포를 한다.