감마 분포의 Special Case

$X \sim Gamma(\alpha = r/2, \beta = 2)$ 인 경우 모수 r 인 카이-자승(Chi-Square) 분포를 따른다고 한다.

$$f(x) = \frac{1}{2^{r/2}\Gamma(r/2)} x^{r/2-1} e^{-x/2}, 0 \le x, 0 < r,$$

NOTATION $Chisquare(r)$,

평균, 분선: $E(X) = r$, $V(X) = 2r$

$X \sim Gamma(\alpha = 1, \beta)$ 인 경우 모수 β 인 지수(exponential) 분포를 따른다고 한다.

$$f(x) = \frac{1}{\beta} e^{-x/\beta}, 0 \le x, 0 < \beta,$$

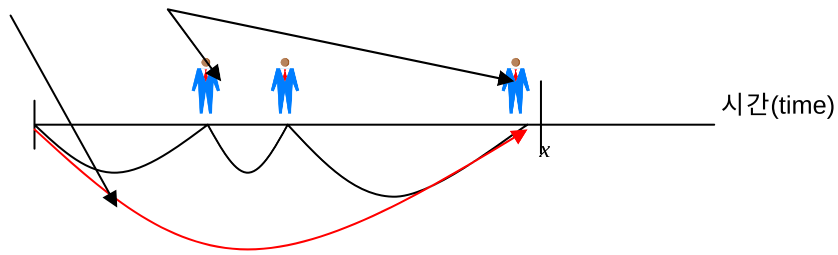
NOTATION $Exponential(\beta)$,

평균, 분산: $E(X) = \beta$, $V(X) = \beta^2$

감마분포와 포아송 분포

만약 $(x_1, x_2, ..., x_\alpha)$ 가 서로 독립이고 모수가 λ 인 지수 분포를 따른다면 $Y = \sum_{i=1}^{\alpha} x_i$ 는 모수가 (α, λ) 인 감마 분포 $Gamma(\alpha, \lambda)$ 따른다. 다음은 포아송 분포와 감마 분포의 관계를 나타낸 것이다. 포아송 분포를 따르는 사건이 일어나는 사이 시간의 분포는 지수분포를 따른다.

$X \sim Gamma(\alpha, \lambda)$, $Y \sim Poisson(x/\lambda)$ ➔ $P(X \le x) = P(Y \ge \alpha)$



이산형 확률변수 X 가 $\sim Poisson(\lambda)$ 을 따른다고 하자. 그리고 연속형 확률변수 Y 을 X 가 일어나는데 걸리는 시간이라 정의하자. 그러면 다음이 성립한다.

$$F(y) = P(Y \le y) = P(\text{사건이}(0, y)\text{구간에서일어남})$$
$$= 1 - P(\text{사건이}(0, y)\text{구간에서일어나지않음}) = 1 - e^{-\lambda y}.$$

그러므로 Y 의 확률밀도함수는 $f(y) = \lambda e^{-\lambda y}, y > 0 \sim Exponential(\beta = 1/\lambda)$ 이다.

감마분포를 따르는 것은 MGF에 의해 보이면 된다.

정리(THEOREM)

$X \sim Gamma(\alpha, \beta)$ 인 경우 $\dfrac{2X}{\beta} \sim Chisquare(r = 2\alpha)$ 이다..

PROOF 6장에서 다룰 것이다.

EXAMPLE 4-11 지수분포의 무기억성

$X \sim Exponential(\beta)$ 이고 a, b 가 양의 상수인 경우 $P(X > a + b \mid X > a) = P(X > b)$ 임을 보이시오, ※무기억성(Memoryless property)이라 한다. (Recall: Geometric dist.)

HOMEWORK #11-1 DUE 4월 26일

$X \sim \exp onential(\beta)$ 이고 확률변수 Y 을 다음과 같이 정의하자.

$$Y = k \quad \text{if} \quad k - 1 \le X < k \quad \text{for} \quad k = 1, 2, \ldots$$

확률 $P(Y = k)$ 을 구하고 이것을 이용하여 이산형 확률변수 Y의 확률밀도함수를 구하시오.

4.6 베타 분포(Beta Dist.)

시간의 비에 대한 분포 (예) 감마분포를 따르는 두 확률변수의 비 $\dfrac{X}{X + Y}$

정의(DEFINITION)

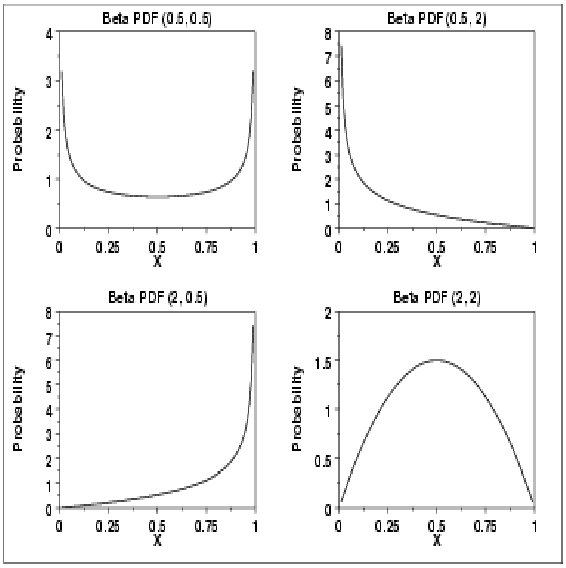
아래 확률밀도함수를 갖는 확률변수 X 는 모수 α, β 을 갖는 베타 분포를 따른다.

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

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

NOTATION $X \sim Beta(\alpha, \beta)$

In SAS: PDF('BETA', x, α, β)



Mean & Variance

$$E(X) = \frac{\alpha}{\alpha+\beta}, \quad V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

PROOF

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

분산에 대한 증명은 숙제로 남겨둔다. Q.E.D.

HOMEWORK #11-2 DUE 4월 26일

$X \sim Beta(\alpha, \beta)$ 일 경우 $V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ 임을 보이시오.

EXAMPLE 4-12 베타분포의 special case

$X \sim Beta(\alpha = 1, \beta = 1)$ 이면 $X \sim Uniform(0,1)$ 을 따른다.

HOMEWORK #11-3 DUE 4월 26일

은행의 업무 시간 중 바쁜 시간의 비율을 확률변수 X 라 정의하자. 확률변수 X 의 확률밀도함수가 $f(x) = cx^2(1-x)^4, 0 \le x \le 1$ 라 주어져 있다.

(1)상수 c 을 구하시오.

(2)바쁜 시간 비율의 기대값 $E(X)$ 을 구하시오.

4.7 MGF (Moment Generating Function)

정의(DEFINITION) $M_X(t) = E(e^{tx}) = \int e^{tx} f(x)dx$

○ $M_x^{(k)}(t=0) = E(X^k)$: 원점에 대한 k-th 적률(k-th moment about origin)

○MGF 유일성(Uniqueness): 적률생삼함수가 같은 확률변수는 동일한 확률분포함수를 갖는다.

EXAMPLE 4-13 적률생성함수의 유일성(이산형)

아래 적률생성함수를 갖는 확률분포함수를 얻으시오.

(a) $M(t) = [(1/3)e^t + (2/3)]^5$. (b) $M(t) = \frac{e^t}{2-e^t}$ (c) $M(t) = e^{2(e^t - 1)}$

EXAMPLE 4-14 지수분포 MGF

$X \sim Exponential(\beta)$ 의 MGF가 $(\frac{1}{1-\beta t})$ 임을 보이시오.

$$m(t) = E(e^{tX}) = \int_0^\infty e^{tx} \frac{1}{\beta} e^{-x/\beta})dx = \frac{1}{\beta}\int_0^\infty e^{-x(1-\beta t)/\beta} dx (Since \beta/(1-\beta t) > 0 \Rightarrow t < 1/\beta)$$

$$= \frac{(\frac{\beta}{1-\beta t})}{\beta} \int_0^\infty \frac{1}{(\frac{\beta}{1-\beta t})} e^{-x(1-\beta t)/\beta} dx = \frac{1}{1-\beta t}.$$

Q.E.D

HOMEWORK #11-4 DUE 4월 26일

$X \sim Exponential(\beta)$ 의 MGF가 $(\frac{1}{1-\beta t})$ 임을 이용하여 $X \sim Gamma(\alpha, \beta)$ 의 MGF가 $(\frac{1}{1-\beta t})^\alpha$ 임을 증명하시오.

EXAMPLE 4-15 정규분포 MGF

$Z \sim Normal(0,1)$ 의 MGF가 $\exp(\frac{t^2}{2})$ 임을 보이시오.

$$M_Z(t) = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}} dz = e^{t^2/2}$$ Q.E.D.

HOMEWORK #11-5 DUE 4월 26일

(a) $M_X(t) = M_{aX+b}(t) = e^{bt} M_X(at)$ 임을 보이시오.

(b)위의 사실을 이용하여 $X \sim Normal(\mu, \sigma^2)$ 의 MGF가 $\exp(\mu t + \frac{\sigma^2 t^2}{2})$ 임을 보이시오.

4.8 Tchebysheff's Theorem

확률변수 X 가 평균 μ, 분산 σ^2 을 갖는 경우 양의 상수 k 에 대하여 다음이 성립한다.

$$P(|X-\mu|<k\sigma) \ge 1 - \frac{1}{k^2} \quad or \quad P(|X-\mu| \ge k\sigma) \le \frac{1}{k^2}$$

PROOF 3장 참고.

EXAMPLE 4-16 Tchebysheff 정리 적용 예제

은행에서 기다리는 시간이 $Gamma(\alpha = 3.1, \beta = 2)$ 을 따른다고 한다. 고객 한 명이 21.5분을 기다리고 "너무 오래 기다렸다고" 항의한다. 적절한 주장인가?

평균, 분산: $\mu = 3.1 \times 2 = 6.2, \sigma^2 = 3.1 \times 2^2 = 12.4$

$\sqrt{12.4}k = (21.5 - 6.2) \Rightarrow 4.32$. Therefore $1/k^2 = 0.053$

$$P(X - 6.2| \ge 15.3) = 0.053(Tchebysheff)$$
$$= 0.0000(Empirical\ Rule)$$

➔ (유의확률) $Gamma(\alpha = 3.1, \beta = 2)$ 상황 하에서는 일어날 가능성이 매우 작으므로 고객의 불평은 합당하다.

4.9 연속형 확률변수 관계

4장에서 증명되지 못한 관계는 5-6장에서 다루기로 한다.