Chapter 4. 연속형 확률변수

확률변수 X 가 가질 수 있는(서로 다른) 값이 유한(finite)이거나 셀 수 있는(countable) 이산형(discrete) 확률변수에 대한 살펴보았다. 여기서는 확률실험 결과 서로 다른 값이 무한히 발생할 수 있는 연속형(continuous) 확률변수에 대해 다룰 것이다. 연속형에서는 아무리 작은 구간을 설정하더라도 적어도 하나 이상의 값이 관측된다. (예) 키, 강우량, 소득, 수명, 경제 지표 등

이산형인 경우 확률밀도함수(확률변수가 갖는 값에 대해 확률을 할당하는 식, 표, 그래프)를 얻는 것은 어렵지 않다. 그러나 아무리 작은 구간이더라도 많은 값들이 관측될 수 있는 연속형의 경우 확률밀도함수를 얻는 것은 불가능하다. 그럼? 관측된 데이터로부터 히스토그램(histogram)을 그리고, 이것이 알려진 확률밀도함수(Gaussian 분포, t-분포) 중 어느 것과 가장 유사한지 판단하여 이론적 분포를 얻게 된다.

왜 우리는 확률밀도함수에 관심을 갖는가?

(1)표본 데이터의 확률밀도함수는 모집단의 확률밀도함수와 동일하다. 그러나 히스토그램에 의존하는 방식으로는 정확한 분포를 아는 것은 불가능하다.

(2)통계량 $h(X_1, X_1, ..., X_n)$ 의 확률밀도함수는 모집단의 모수(parameter)에 대한 추론에 핵심적인 역할을 한다. 표본의 분포를 알려면(물론 표본 평균은 CLT가 있기는 하지만) 모집단의 분포를 알아야 하는데 일반적으로는 불가능하다. 그러므로 데이터의 특성에 따라 모집단의 분포를 가정하게 된다.

4.1 분포함수 (distribution function)

정의(DEFINITION)

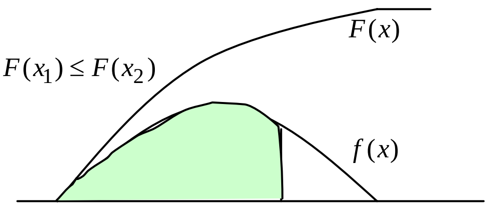
연속형 확률변수 X 의 확률분포함수(probability distribution function) 혹은 누적확률밀도함수(cumulative density function) $F(x)$ (기호)는 $F(x) = P(X \le x),\ for\ -\infty < x < \infty$ 라 정의한다.

$F(x)$ 의 성질

⑴ $F(-\infty) = 0, F(\infty) = 1$

⑵ $F(x)$ 는 비감소(non-decreasing) 함수이다. $If\ x_1 < x_2,\ \text{then}\ F(x_1) \le F(x_2)$

초록색 부분이 $F(x)$ 이다.

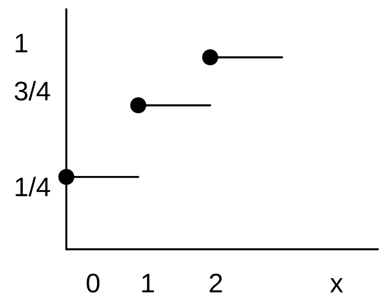


EXAMPLE 4-1 분포함수 구하기

이산형 확률변수 $X \sim Binomial(n=2, p=0.5)$ 에 대하여 분포함수 $F(x)$ 을 얻으시오.

$$F(x) = \begin{cases} 0, x<0 \\ 1/4, 0 \le x < 1 \\ 3/4, 1 \le x < 2 \\ 1, x \ge 2 \end{cases}$$

(이산형 확률변수의 $F(x)$ 는 step 함수)



정의**(DEFINITION)**

⑴만약 $F(x)$ 가 연속이면 확률변수 X 는 연속형이다.

⑵연속형 확률변수의 분포함수를 $F(x)$ 라 하면 확률밀도함수 $f(x)$ 는 다음에 의해 계산 된다. $f(x) = \frac{d}{dx}F(x)$

$f(x)$ 는 $F(x)$ 으로부터 얻어지므로 아래 성질을 갖는다.

연속형 확률밀도함수의 성질

⑴ $f(x) \ge 0$ for any value of x

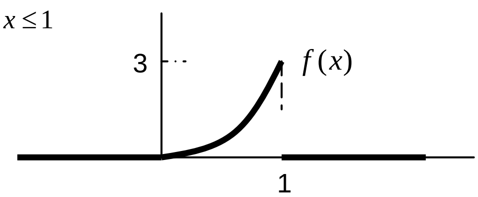
⑵ $\int_{-\infty}^{\infty} f(x)dx = 1$

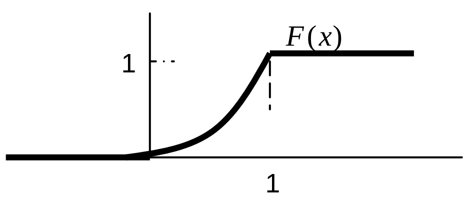
⑶ $P(a \le X \le b) = F(b) - F(a)$

EXAMPLE 4-2 분포함수 구하기(2)

확률변수 X 의 확률밀도함수 $f(x) = 3x^2, 0 \le x \le 1$ 일 경우 분포함수 $F(x)$ 을 구하고 그래프를 그리시오.

$$F(x) = \begin{cases} 0, x<0 \\ x^3, 0 \le x \le 1 \\ 1, x \ge 1 \end{cases}$$





EXAMPLE 4-3 분포함수 이용하기

확률변수 X 의 확률밀도함수가 $f(x) = cx^2, 0 \le x \le 2$ 일 경우

(1)상수 c 을 구하시오. (2)분포함수 $F(x)$ 구하시오. (3) $P(1 < X \le 2)$ 을 계산하시오.

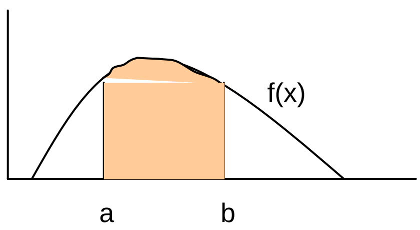
$c = 3/8$

$F(x) = x^3/8$, $0 \le x \le 2$

$7/8$

확률과 분포함수

$$P(a \le X \le b) = F(b) - F(a)$$



HOMEWORK #9-1

확률변수 X 의 확률밀도함수가 $f(x) = cx, 0 \le x \le 2$ 일 경우.

(1)상수 c 을 구하시오.

(2)확률분포함수 $F(x)$ 을 구하시오.

(3) $F(x)$ 이용하여 $P(1 < X \le 2)$ 계산하시오.

HOMEWORK #9-2

확률변수 X 의 분포함수는 $F(x) = \begin{cases} 0, x \le 0 \\ x/8, 0 < x < 2 \\ x^2/16, 2 \le x < 4 \\ 1, x \ge 4 \end{cases}$ 이다.

(1)확률밀도함수 $f(x)$ 을 구하시오.

(2) $P(X > 1.5)$ 계산하시오.

(3) $P(X \ge 1 | X \le 3)$ 계산하시오.

4.2 기대값(Expected value)

□평균 mean($\mu, \bar{x}$): 데이터의 중앙 위치 모수 $E(X)$

□분산 variance(σ^2, s^2): 산포(흩어짐) 모수 $E(X-E(X))^2 = E(X^2) - E(X)^2$

□표준편차 standard deviation(σ, s): 분산의 제곱근

연속형 확률변수의 확률밀도함수(pdf)를 얻는 것은 불가능하지만, 적률($E(X^k)$)과 경험적 법칙(Empirical Rule, $\pm 2\sigma$, 95%)이나 Tchebysheff's inequality($P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$)을 이용하여 확률변수(데이터)의 분포를 판단할 수 있다.

정의(DEFINITION)

□연속형 확률변수 X의 기대값(expected value)은 $E(X) = \int xf(x)dx$으로 정의한다.

□확률변수 X의 함수 $g(X)$의 기대값은 $E(g(X)) = \int g(x)f(x)dx$로 정의된다.

□만약 $g(X) = (X - E(X))^2$이면, $g(X)$의 기대치는 확률변수 X의 분산이다.

정리(THEOREM)

(1)상수 c에 대하여 $E(c) = c$이 성립한다.

$(2) E[cg(X)] = cE[g(X)]$

$(3) E[g_1(X) + g_2(X) + ... + g_k(X)] = E[g_1(X)] + E]g_2(X)] + ... + E]g_k(X)]$

PROOF obvious

HOMEWORK #9-3

연속형 확률변수 X는 평균 μ이고 분산 σ^2을 갖는다. 상수 a, b에 대하여 다음이 성립함을 보이시오. $E(aX+b) = a\mu + b$ and $V(aX+b) = a^2\sigma^2$

EXAMPLE 4-4 평균과 분산 구하기

만약 $X \sim f(x) = 1/2, 59 \leq x \leq 61$이면, 확률변수 X의 평균과 분산을 구하시오.

60 / 0.333

EXAMPLE 4-5 평균과 분산 구하기

만약 $X \sim f(x) = (3/2)x^2, 0 \le x \le 1$ 이면, 확률변수 $W = (5 - 0.5X)$ 의 평균과 분산을 구하시오.

4.8 / 0.039

HOMEWORK #9-4

확률밀도함수 $f(x) = 2x, 0 \le x \le 1$ 을 갖는 확률변수 X 에 대하여.

(1) $E(X), V(X)$ 을 계산하시오.

(2)확률변수 $W = 200X - 60$ 의 평균과 분산을 구하시오..

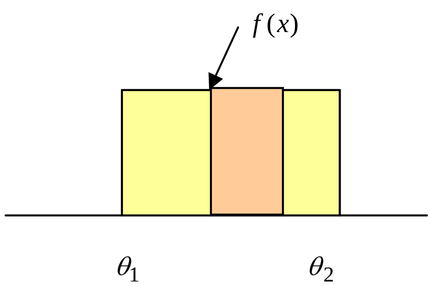
4.3 균일분포 (Uniform Dist.)

버스는 오전 8:00와 8:10 사이에 반드시 정차하고 10분 사이 어느 구간에서든 버스가 도착하는 가능성은 동일하다고 하자. 즉 임의의 구간에 버스가 도착할 확률은 시간 길이에 비례한다.

정의(DEFINITION)

확률변수 X 을 임의의 구간(θ_1, θ_2) 사이의 시간이나 거리라 하자. 만약 X 의 pdf가 다음과 같다면 확률변수 X 는 구간 (θ_1, θ_2) 에서 연속형 균일분포를 갖는다고 한다.

$$f(x) = \frac{1}{\theta_2 - \theta_1}, \theta_1 \le x \le \theta_2$$



$\boxed{\text{NOTATION}}$ $X \sim Uniform(\theta_1, \theta_2)$

In SAS: PDF('UNIFORM', x, θ_1, θ_2)

$\theta_1 = 0, \theta_2 = 1$ 인 균일분포는 난수(random number) 생성에 이용된다.

평균과 분산

$$E(X)=\frac{\theta_1+\theta_2}{2},\quad V(X)=\frac{(\theta_2-\theta_1)^2}{12}\ (\text{숙제})$$

EXAMPLE 4-6 균일분포 확률

임의의 30분 동안 은행에 고객이 방문하는 시간은 균일 분포를 따른다고 하자. 마지막 5분에 고객이 도착할 확률을 계산하시오.

1/6

EXAMPLE 4-7 균일분포 확률(2)

확률변수 $X \sim Uniform(50,70)$ 라면 $P(X \geq 65 \mid X \geq 55)$ 을 구하시오.

1/3

HOMEWORK #10-1

$X \sim Uniform(\theta_1,\theta_2)$ 의 평균과 분산이 각각 $E(X)=\frac{\theta_1+\theta_2}{2}$, $V(X)=\frac{(\theta_2-\theta_1)^2}{12}$ 임을 보이시오.

HOMEWORK #10-2

만약 낙하산이 marker A와 B 사이에 임의의 지점에 떨어진다고 하자.

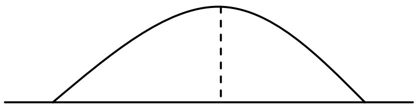
(1)낙하산이 B 보다 A 지점에 더 가까이 떨어질 확률을 구하시오.

(2)낙하산 떨어진 지점에서 지점 A까지의 거리가 지점 B까지의 거리의 3배 이상일 확률을 구하시오.

(3)3개의 낙하산 중 정확하게 한 개만 지점 B에 가까이 떨어질 확률을 구하시오.

4.4 정규분포(Normal Dist.)

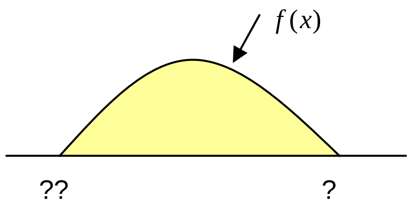
종모양(Bell-shaped), 경험적 법칙(Empirical Rule), 측정 오차(Measurement error)



정의(DEFINITION)

확률변수 X 의 확률밀도함수가 아래와 같으면 평균 μ, 분산 σ^2 을 갖는 정규분포를 따른다고 한다.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty \le x \le \infty$$



$\boxed{\text{NOTATION}}$ $X \sim Normal(\mu, \sigma^2)$

In SAS: PDF('NORMAL', x, μ, σ)

평균과 분산

$E(X) = \mu$, $V(X) = \sigma^2$: 증명은 MGF을 구한 후 그것을 이용하여 보일 것이다. (skip now)

표준정규분포(Standard Normal dist.)

확률변수 $X \sim Normal(\mu, \sigma^2)$ 일 경우 $Z = \dfrac{X-\mu}{\sigma}$ (표준화: Standardization)은 평균이 0이고 분산이 1인 정규분포함수를 따른다. $\mu = 0, \sigma^2 = 1$ 이를 표준정규분포라 한다.

$$Z = \frac{X-\mu}{\sigma} \sim Normal(0,1)$$

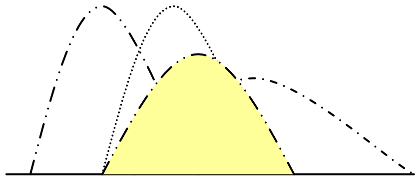
정규분포에 대한 증명은 MGF 함수를 구한 후 다루기로 한다. 평균 0, 분산 1은 기대값의 성질에 의해 증명된다.

EXAMPLE 4-8 평균/분산 구하기

$X \sim ?(\mu, \sigma^2)$ (어떤 분포함수라도)일 경우 $Z = \dfrac{X-\mu}{\sigma}$ 의 평균과 분산을 구하시오.

표준정규분포표

정규분포의 확률은 평균과 분산에 따라 다르다. 와~ 엄청나게 많은 분포표가 필요하다? 해결방법은? 표준정규분포를 이용한다.



EXAMPLE 4-9 정규분포 확률계산하기

확률변수 X 가 표준정규분포를 따른다. $X \sim Normal(0,1)$ 다음 확률을 계산하시오.

(1) $P(X > 2)$ (2) $P(-2 \le X \le 2)$ (3) $P(X \le 1.73)$

0.0228 / 0.9544 / 0.4582

EXAMPLE 4-10 정규분포 확률계산하기

학생들의 SAT 점수는 평균 75, 표준편차 10인 정규분포를 따른다고 하자. 점수 80~90 사이의 학생 비율은 얼마나 되나? 60점~80점 사이 학생의 비율은?

상위 10% 학생의 점수는 몇 점인가?

0.2417 / /

HOMEWORK #10-3

확률변수 X 가 표준정규분포를 따른다. $X \sim Normal(0,1)$ 다음 확률을 계산하시오.

(1) $P(-0.9 \le X < 0)$ (2) $P(-1.56 \le X \le 2)$ (3) $P(0.3 < X \le 1.56)$

HOMEWORK #10-4

확률변수 X 가 표준정규분포를 따른다. $X \sim Normal(0,1)$ 다음 상수를 구하시오.

(1) $P(k \le X) = 0.8643$ (2) $P(-k < X < k) = 0.9$ (3) $P(-k < X < k) = 0.99$

HOMEWORK #10-5

○○회사 생산하는 볼트 지름의 크기는 평균 950 millimeters, 표준편차 10 millimeters인 정규분포를 따른다고 한다.

(1)볼트를 하나 선택했을 때 그것의 지름이 947~958 millimeters일 확률을 계산하시오.

(2)볼트의 지름이 상수 C 보다 적을 확률이 0.8531일 경우 상수 C 을 구하시오.

4.5 감마 분포 (Gamma Dist.)

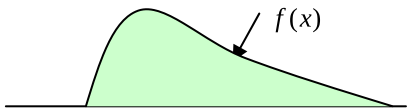
우로 치우친 분포(skewed to the right, positively skewed), 기다리는 시간, 수명

정의(DEFINITION)

다음 확률밀도함수를 갖는 확률변수는 감마분포를 따른다고 한다.

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

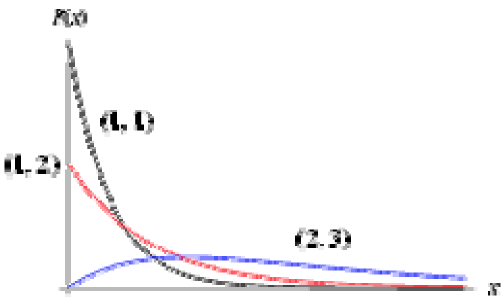
$$where\ \Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}e^{-x}dx$$

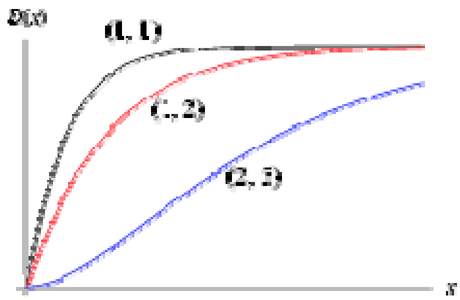


NOTATION $X \sim Gamma(\alpha,\beta)$

In SAS: PDF('GAMMA', x,α,β)

α 는 형태(shape) 모수, β 는 크기(scale) 모수이다. 다음 그래프는 모수에 따른 감마 확률밀도함수를 그린 것이다.





평균과 분산

$E(X)=\alpha\beta$, $V(X)=\alpha\beta^2$

PROOF 확률밀도함수의 전체 적분은 1이므로 $\int_0^\infty x^{\alpha-1}e^{-x/\beta}=\beta^\alpha\Gamma(\alpha)$ 이다.

$$E(X)=\int_0^\infty x\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}dx$$
$$=\frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty x^\alpha e^{-x/\beta}dx=\frac{\beta^{\alpha+1}\Gamma(\alpha+1)}{\beta^\alpha\Gamma(\alpha)}=\alpha\beta$$

같은 방법으로 $E(X^2)=\dfrac{\beta^{\alpha+1}\Gamma(\alpha+2)}{\beta^\alpha\Gamma(\alpha)}=\alpha(\alpha+1)\beta^2$.

그러므로 $V(X)=E(X^2)-[E(X)]^2=\alpha\beta^2$.

감마함수의 성질 $\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx$

○ α 가 정수인 경우 $\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$.

PROOF 부분 적분(Integral by parts)이나 표 적분을 이용하자.

부분 적분: $d(uv)=(du)v+u(dv)\Rightarrow\int d(uv)=\int v(du)+\int u(dv)$
$\Rightarrow\int u(dv)=uv-\int v(du)$

$u=x^{\alpha-1}$, $dv=e^{-x}dx$ 라 놓자. ➔ $du=(\alpha-1)x^{\alpha-2}dx$, $\int dv=\int e^{-x}dx\Rightarrow v=-e^{-x}$

$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx=-x^{\alpha-1}(e^{-x})|_0^\infty-\int-e^{-x}(\alpha-1)x^{\alpha-2}dx$ 이다.

그러므로 $\Gamma(\alpha)=(\alpha-1)\int e^{-x}x^{\alpha-2}=\int_0^\infty x^{\alpha-1}e^{-x}dx$. *Q.E.D.*

○ α 가 정수인 경우 $\Gamma(\alpha)=(\alpha-1)!$.

PROOF 위의 결과와 $\Gamma(1)=1$ 을 이용한다.

○ $\Gamma(1/2)=\sqrt{\pi}$:

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.

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

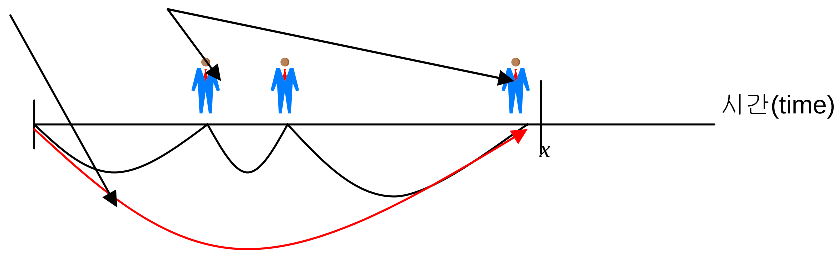
NOTATION $Exponential(\beta)$,

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 $\dfrac{X}{X+Y}$

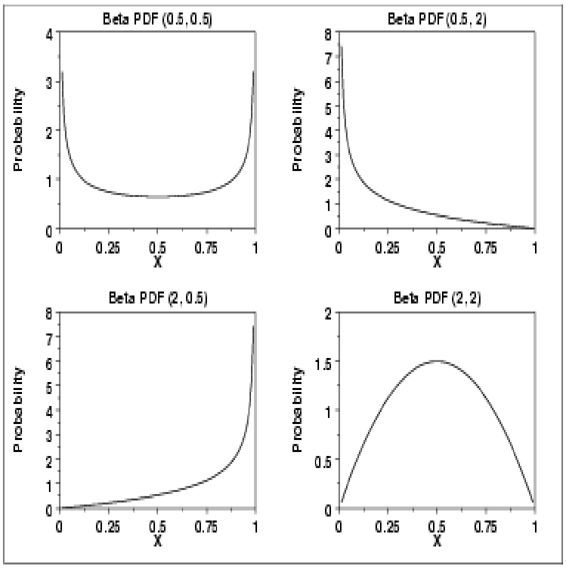
정의(DEFINITION)

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 확률 밀도함수가 $f(x) = cx^2(1-x)^4, 0 \le x \le 1$ 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) $M(t) = \dfrac{e^t}{2-e^t}$ (c) $M(t) = e^{2(e^t-1)}$

EXAMPLE 4-14 지수분포 MGF

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

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

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

$\boxed{Q.E.D}$

HOMEWORK #11-4 DUE 4월 26일

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

EXAMPLE 4-15 정규분포 MGF

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

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

$\boxed{\text{Q.E.D.}}$

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.

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

$\boxed{\text{PROOF}}$ 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 $1/k^2 = 0.053$

$$\begin{aligned} P(|X-6.2| \ge 15.3) &= 0.053(Tchebysheff) \\ &= 0.0000(Empirical\ Rule) \end{aligned}$$

➔ (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.