표본 평균

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

표본 분산

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

표본 표준편차

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

$$y_i = c x_i$$

$$\rightarrow \bar{y} = c\bar{x} \qquad s_y^2 = c^2 s_x^2$$

조건부 확률.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

독립사건

$$P(A \cap B) = P(A) \cdot P(B)$$

베르누이

$$p(x) \begin{cases} 1-\alpha & x=0 \\ \alpha & x=1 \\ 0 & \text{otherwise} \end{cases}$$

기하분포.

$$p(x) \begin{cases} (1-p)^{x-1} p & x = 1, 2, \cdots \\ 0 & \text{otherwise} \end{cases}$$

CDF

$$F(x) = P(X \le x)$$

E(X)

$\mu_x = \sum x p(x)$ (이산확률분포)

$E(X) = p$ (베르누이)

$E(X) = 1/p$ (기하분포)

$$E(h(x)) = \sum h(x) p(x)$$

$$\therefore E(aX+b) = aE(X) + b$$

$$V(X) = \sum (x-\mu)^2 p(x) = E((X-\mu)^2)$$

$$= E(X^2) - E(X)^2$$

$$V(aX+b) = a^2\sigma^2 \;/\; \sigma(aX+b) = |a|\sigma$$

$V(X) = \frac{1-p}{p^2}$ (기하분포)

이항확률분포.

$$p(x) \begin{cases} {}_nC_x \, p^x (1-p)^{n-x} & x = 0, 1, 2, \cdots, n \\ 0 & \text{otherwise.} \end{cases}$$

$E(X) = np$

$V(X) = np(1-p)$

$\sigma(X) = \sqrt{np(1-p)}$

포아송분포.

$$p(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \quad \begin{cases} x = 1, 2, \cdots \\ \lambda > 0 \end{cases}$$

$E(X) = \lambda$

$V(X) = \lambda$

이항분포의 포아송분포 근사.

$b(x; n, p) \rightsquigarrow p(x; \lambda)$ $\quad \lambda = np.$ with $\begin{cases} \text{large } n \\ \text{small } p \end{cases}$

독립인 두 포아송분포의 합.

$\lambda = \lambda_1 + \lambda_2$ & $Z = X + Y$

$$P_Z(z) = P(Z = z) = \frac{e^{-\lambda} \cdot \lambda^z}{z!}$$ $\rightsquigarrow$ X+Y follow poisson!

포아송프로세스.

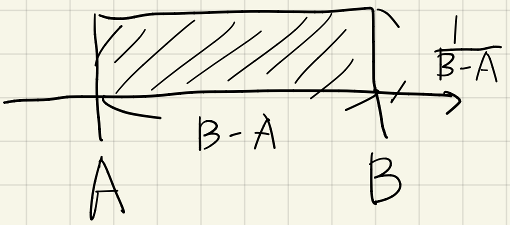
$\lambda \cdot \Delta t$ $\quad$ t 기간 사이 발생하는 사건

$$\sim P_k(t) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}$$

연속확률분포.

pdf $\quad p(a \le x \le b) = \int_a^b f(x)\,dx$, $\quad f(x) \ge 0$. $\quad \int_{-\infty}^{\infty} f(x)\,dx = 1$

uniform distribution



cdf $\quad \int_{-\infty}^{x} f(y)\,dy = P(X \le x) = F(X)$

$F'(x) = f(x) \qquad P(a \le x \le b) = F(b) - F(a)$

$\qquad\qquad\qquad P(x \ge a) = 1 - F(a)$

$100p^{th}$ percentile

$p = F(\eta(p)) = P(X \le \eta(p))$, $\quad \tilde{\mu} = \text{median} = 0.5$.

$$E(X) = \int_{-\infty}^{\infty} x\,p(x)\,dx = \mu_x$$

$E(h(x)) = \int_{-\infty}^{\infty} h(x)\,p(x)\,dx$, $\quad E(aX+b) = aE(X) + b$.

$$V(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\,dx \qquad \sigma_X = \sqrt{V(X)}$$

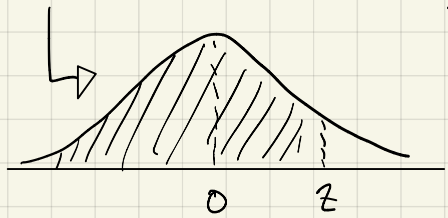
$$= \int_{-\infty}^{\infty} x^2 p(x)\,dx - \mu^2 = E(X^2) - E(X)^2$$

정규분포 .

$X \sim N(\mu, \sigma^2)$ $\quad E(X) = \mu$, $\quad V(X) = \sigma^2$

표준정규분포 -

$X \sim N(0,1)$ ∴ Z $\qquad \Phi(z) = P(Z < z)$



표준화. $\quad Z = \dfrac{X-\mu}{\sigma}$

$100p^{th}$ percentile $\rightarrow$ $\mu + (100p^{th}$ percentile $N(0,1)) \times \sigma$

이항분포의 정규분포 근사 . $\quad b(x; n, p)$

$\mu = np$. $\quad \sigma = \sqrt{np(1-p)}$

Binom $P(X \le x)$

$\approx$ Normal $P(X \le x+0.5)$

$= \Phi\left(\dfrac{x+0.5-np}{\sqrt{np(1-p)}}\right)$

감마분포. $\Gamma(\alpha) = (\alpha-1)!$ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} = \frac{1}{\beta^\alpha (\alpha-1)!} \times x^{\alpha-1} \times e^{-x/\beta} & (x \geq 0) \\ 0 & \text{otherwise.} \end{cases}$$

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

지수분포 $\alpha = 1$, $\beta = \frac{1}{\lambda}$ 인 감마분포.

pdf
$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & \text{otherwise.} \end{cases}$$

$E(X) = \frac{1}{\lambda}$ $V(X) = \frac{1}{\lambda^2}$

cdf
$$F(x;\lambda) \begin{cases} 1-e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

$F(X) = P(X \leq x)$

카이제곱분포 $\alpha = \frac{r}{2}$ $\beta = 2$ 인 감마분포.

$r = df$ (자유도)

$$f(x;r) = \frac{1}{\beta^{\frac{a}{2}} (\frac{r}{2}-1)!} \times x^{\frac{r}{2}-1} \times e^{-x/2}$$

Joint Probability Mass Function

$$p(x,y) = P(X=x \;\&\; Y=y)$$

marginal
$$P_x(X) = \Sigma_y\, p(x,y)$$
$$P_y(Y) = \Sigma_x\, p(x,y)$$

Joint Probability Density Function

$a \le x \le b \qquad c \le y \le d$

$$\int_a^b \int_c^d f(x,y)\,dy\,dx = P(X,Y)$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx = 1$$

marginal
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

독립확률변수. $X \,\&\, Y$

$$\begin{cases} p(x,y) = P_X(x) \cdot P_Y(y) \\ f(x,y) = f_X(x) \times f_Y(y) \end{cases}$$

$\boxed{\text{cf}}$ 232p 233p

조건부분포.

Y의 확률밀도, $X=x$일때, $f_{Y|X}(y|x) = \dfrac{f(x,y)}{f_X(x)}$ $\begin{cases} -\infty \le y \le \infty \\ f_X(x) > 0 \end{cases}$

$$P_{Y|X}(y|x) = \frac{P(x,y)}{P_X(x)} \quad \{ \ P_X(x) > 0$$

$E(X)$ of joint

$$E(h(x,y)) \begin{cases} \sum_x \sum_y h(x,y)\, p(x,y) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y)\, p(x,y)\, dx\, dy \end{cases}$$

$$Cov(X,Y) = E((X-\mu_x)(Y-\mu_y))$$

$$\begin{cases} \sum_x \sum_y (x-\mu_x)(y-\mu_y)\, p(x,y) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y)\, f(x,y)\, dy\, dx \end{cases}$$

$$= E(XY) - E(X)E(Y)$$

상관계수

$$\rho_{x,y} = \frac{Cov(x,y)}{\sigma_x \sigma_y} \qquad -1 \le \rho_{xy} \le 1$$

X & Y $\rightsquigarrow$ independent, $\rho = 0$.

표본평균의 특징

$E(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$ (독립여부 상관X)

$X_1, X_2, \cdots, X_n$ 이 독립 $\rightsquigarrow Cov(X,Y) = 0$.

$\rightsquigarrow V(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2$

$X_1, X_2, \cdots, X_n$ 이 독립 X

$\rightsquigarrow V(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j Cov(X_i, X_j)$

$X_1, X_2, \cdots, X_n$ 이 임의 선택된 표본 & $\overset{(mean)}{\mu}, \overset{(sd)}{\sigma}$

$E(\bar{X}) = \mu \qquad V(\bar{X}) = \frac{\sigma^2}{n} \qquad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

$T_n = X_1 + X_2 + \cdots + X_n$ 이 표본전체,

$E(T_n) = n\mu \qquad V(T_n) = n\sigma^2 \qquad \sigma(T_n) = \sqrt{n}\sigma$

중심극한정리.

$X_1, X_2, \cdots, X_n \Rightarrow$ 임의표본. mean : μ. sd : σ

n이 충분히 크면, $\mu_{\bar{X}} = \mu$. $\sigma_{\bar{X}}^2 = \left(\frac{\sigma}{\sqrt{n}}\right)^2 = \frac{\sigma^2}{n}$

$$\therefore \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \quad n \to \infty$$

확률변수의 차

$Y = X_1 - X_2$

$E(Y) = E(X_1) - E(X_2)$ $\quad V(Y) = V(X_1 - X_2) = 1^2 V(X_1) + (-1)^2 V(X_2)$

$= V(X_1) + V(X_2) = \sigma_1^2 + \sigma_2^2$

$\sigma_Y^2 = \sigma_1^2 + \sigma_2^2 \quad \sigma_Y = \sqrt{\sigma_1^2 + \sigma_2^2}$

$$\int f \cdot g' = fg - \int f'g\,dx$$