예) X는 어느 공장에서 하루동안 생산되는 제품중 불량품의 갯수, Y는 생산라인 1에서의 하루동안 생산되는 제품중 불량품의 개수. 이 공장에서 하루동안 생산되는 제품중 불량품의 갯수가 *n*개일때, 불량품중 생산라인 1에서 생산된 제품의 개수의 분포는?

독아 송

이항분포

$Y \sim Pois(\lambda_1)$

Z: 나머지생산라인 불량품의 개수. $Z \sim Pois(\lambda_2)$.

$X = Y + Z$, Y, Z: 독립 $\Rightarrow X \sim Pois(\lambda_1 + \lambda_2)$

$$P(Y=x \mid X=n) = \frac{P(Y=x, X=n)}{P(X=n)} = \frac{P(Y=x, Z=n-x)}{P(X=n)} = \frac{P(Y=x)\cdot P(Z=n-x)}{P(X=n)}$$

$$= \frac{\cancel{e^{-\lambda_1}}\cdot\frac{\lambda_1^x}{x!}\cancel{e^{-\lambda_2}}\cdot\frac{\lambda_2^{n-x}}{(n-x)!}}{\cancel{e^{-(\lambda_1+\lambda_2)}}\cdot\frac{(\lambda_1+\lambda_2)^n}{n!}}$$

$$= \frac{n!}{x!(n-x)!}\cdot\frac{\lambda_1^x\cdot\lambda_2^{n-x}}{(\lambda_1+\lambda_2)^n}$$

$$= {}_nC_x\cdot\frac{\lambda_1^x\cdot\lambda_2^{n-x}}{(\lambda_1+\lambda_2)^x\cdot(\lambda_1+\lambda_2)^{n-x}}$$

$$= {}_nC_x\cdot\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x\cdot\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-x}$$

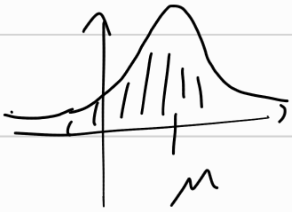
$$= {}_nC_x\cdot\left(\boxed{\frac{\lambda_1}{\lambda_1+\lambda_2}}\right)^x\cdot\left(1-\boxed{\frac{\lambda_1}{\lambda_1+\lambda_2}}\right)^{n-x}$$

$$\left(B\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)\right)$$

$B(n,p)$

${}_nC_x\cdot p^x(1-p)^{n-x}$

$\boxed{N(\mu, \sigma^2)}$



$\boxed{\frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}(x-\mu)^2}}$

$E(X), Var(X) \Rightarrow E(X) E(Y) \; Var(X) \; Var(Y), Cov(X,Y)$

$, Corr(X,Y) \ldots$

$$\begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}, \underbrace{\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \vdots & & & \\ \sigma_{k1} & & \cdots & \sigma_{kk} \end{pmatrix}}_{} \right)$$

$= \Sigma$

$\sigma_{ij} = Cov(X_i, X_j)$

$\boxed{\frac{1}{2}(x-\mu)\sigma^{-2}(x-\mu)}$

$$\text{pdf } f(x) = (2\pi)^{-\frac{k}{2}} \cdot |\Sigma|^{-\frac{1}{2}} \cdot e^{-\frac{1}{2}(X-\mu)^t \Sigma^{-1} (X-\mu)}$$

행렬식

k-차원 벡터

k×k 행렬

$$\begin{pmatrix} \mathbb{X}_1 \\ \hline \mathbb{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \hline \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$$

$\mathbb{X}_1 \rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbb{X}_2 \rightarrow \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix}$

Σ_{11}: 2x2 행렬, $Var(\mathbb{X}_1)$

Σ_{12}: $Cov(\mathbb{X}_1, \mathbb{X}_2)$ 2x3 행렬

Σ_{21}: 3x2 행렬 $Cov(\mathbb{X}_2, \mathbb{X}_1)$

Σ_{22}: $Var(\mathbb{X}_2)$ 3x3 행렬

$$\Sigma_{11} = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var(X_2) \end{pmatrix}$$

$$\boxed{\mathbb{X}_1 \mid \mathbb{X}_2 = \mathbb{x}} \sim N(\bigcirc, \bigcirc)$$

문제 1. 다음과 같은 분포를 가지는 무한모집단에서의 크기 2의 랜덤표본을 X_1, X_2라 하자.

x	0	1	2
$P(X=x)$	0.2	0.4	0.4

(a) X_1의 기댓값 $E(X_1)$와 분산 $Var(X_1)$를 구하여라.

(b) 표본평균 $\overline{X} = \dfrac{X_1+X_2}{2}$의 확률분포를 구하여라.

(c) $E(\overline{X})$와 $Var(\overline{X})$를 계산하고 $E(X)$, $Var(X)$ 와의 관계에 대해 설명하여라.

$\boxed{1, 2} \rightarrow$

(검정, 회귀분석
EM, HMM, 나이브베이즈)

(a) $E(X) = 0 \times 0.2 + 1 \times 0.4 + 2 \times 0.4 = 1.2$

$E(X^2) = 0 + 1^2 \times 0.4 + 2^2 \times 0.4 = 2 \Rightarrow Var(X) = 2 - 1.2^2 = 0.56$

(b) $\boxed{\overline{X} = \dfrac{X_1+X_2}{2}}$

	0	0.5	1	1.5	2
	0.04	0.16	0.32	0.32	0.16

$P(\overline{X}=0) = P(X_1=0)P(X_2=0) = 0.2 \times 0.2 = 0.04$

$P(\overline{X}=0.5) = P(X_1=1)P(X_2=0) + P(X_1=0)P(X_2=1) = 0.2 \times 0.4 + 0.4 \times 0.2 = 0.16$

$P(\overline{X}=1) = \underset{(0,2)(2,0)(1,1)}{\cdots} = 0.08 + 0.08 + 0.16 = 0.32$

$P(\overline{X}=1.5) = \underset{(1,2)(2,1)}{} = 0.4 \times 0.4 \times 2 = 0.32$

$P(\overline{X}=2) = \underset{(2,2)}{} = 0.4 \times 0.4 = 0.16$

(c) $E(\overline{X}) = E(X_1)$, $Var(\overline{X}) = \frac{1}{n} Var(X_1)$

$E(\overline{X}) = 0.5 \times 0.16 + 1 \times 0.32 + 1.5 \times 0.32 + 2 \times 0.16 = \cdots = 1.2 = E(X_1)$

$Var(\overline{X}) = \cdots = 0.28 = \frac{1}{2} \cdot Var(X_1)$

문제 2. 유한모집단 $\{1, 3, 3, 4\}$에서 크기 2인 표본을 단순랜덤비복원추출하였을 때, 표본평균의 확률분포를 구하여라.

$= \{1, 3_a, 3_b, 4\}$

	2	2.5	3	3.5
	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$

$\bar{X}\left(=\frac{X_1+X_2}{2}\right)=2$: $(X_1, X_2) = (1, 3)$ ✓ $(1, 3_a), (1, 3_b)$

전체 경우의 수 : ${}_4C_2 = 6$

$\bar{X} = 2.5$: $(X_1, X_2) = (1, 4)$ ✓ $(1, 4)$

$\bar{X} = 3$: $(3, 3)$ ✓ $(3_a, 3_b)$

$\bar{X} = 3.5$: $(3, 4)$ ✓ $(3_a, 4), (3_b, 4)$