M368 9/14

$\vec{X} \sim \text{Multin}_K(n, \vec{p})$

K ← dimension

$K=2 \quad \vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$

$\text{Deg}(n-x_2) = P_{X_1|X_2}(x_1, x_2) := P(X_1 = x_1 | X_2 = x_2) = \frac{P(x_1, x_2)}{P(x_2)}$

$P(x_2) = \text{Bin}(n, p_2) = \text{Bin}(n, 1-p_1)$

$J_n := \{0, \dots, n\}$

$= \frac{\binom{n}{x_1, x_2} p_1^{x_1} p_2^{x_2}}{\binom{n}{x_2} p_2^{x_2} (1-p_2)^{n-x_2}} = \frac{\frac{n!}{x_1! \cancel{x_2!}} \mathbb{1}_{x_1+x_2=n} \mathbb{1}_{x_1 \in J_n} \cancel{\mathbb{1}_{x_2 \in J_n}} p_1^{x_1} p_2^{x_2}}{\frac{n!}{\cancel{x_2!}(n-x_2)!} \cancel{\mathbb{1}_{x_2 \in J_n}} p_2^{x_2} p_1^{n-x_2}}$ $\quad x_1 = n - x_2 \quad \mathbb{1}^u_{x_2 \in J_n}$

Define a ratio of indicators: $\mathbb{1}^u_A := \frac{\mathbb{1}_A}{\mathbb{1}_A} = \begin{cases} 1 & \text{if } A \\ \text{undefined} & \text{if } A^c \end{cases}$

$= \underbrace{\frac{(n-x_2)!}{x_1!}}_{=1 \text{ when } x_1 = n-x_2} \mathbb{1}_{x_1 = n-x_2} \underbrace{\mathbb{1}_{x_1 \in J_n}}_{=1 \text{ when } x_1 = n-x_2 \text{ and } x_2 \in J_n} \mathbb{1}^u_{x_2 \in J_n} = \text{Deg}(n-x_2) \mathbb{1}^u_{x_2 \in J_n}$

$P(A|B) = \frac{P(A \cap B)}{P(B)}$ undef $P(B)=0$

$\begin{bmatrix} p_1 \\ \vdots \\ p_K \end{bmatrix} \quad \sum \vec{p}_i = 1$

Let's generalize the result a bit... $\vec{X} \sim \text{Multin}_K(n, \vec{p})$

$P_{\vec{X}_{-j}|X_j}(\vec{x}_{-j}, x_j) = P_{\vec{X}}(\vec{x}) / P_{X_j}(x_j)$ all elements of vector rv X except j'th component

$\sim \text{Multin}_{K-1}(n - x_j, ?)$

$= \frac{\text{Multin}_K(n, \vec{p})}{\text{Bin}(n, p_j)} = \frac{\binom{n}{x_1, \dots, x_K} p_1^{x_1} \cdots p_{j-1}^{x_{j-1}} \cancel{p_j^{x_j}} p_{j+1}^{x_{j+1}} \cdots p_K^{x_K}}{\binom{n}{x_j} \cancel{p_j^{x_j}} (1-p_j)^{n-x_j}}$

$= \frac{\frac{n!}{x_1! \cdots \cancel{x_j!} \cdots x_K!} \cdot \overbrace{\mathbb{1}_{x_1 + \dots + x_j + \dots + x_K = n}}^{x_1 + \dots + x_{j-1} + x_{j+1} + x_K = n - x_j} \mathbb{1}_{x_1 \in J_n} \cdots \cancel{\mathbb{1}_{x_j \in J_n}} \cdots \mathbb{1}_{x_K \in J_n} p_1^{x_1} \cdots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \cdots p_K^{x_K}}{\frac{n!}{\cancel{x_j!}(n-x_j)!} \cancel{\mathbb{1}_{x_j \in J_n}} (1-p_j)^{n-x_j}}$ $\quad \mathbb{1}^u_{x_j \in J_n}$

let $n' = n - x_j$, <u>note</u> $x_1 + \dots + x_K = n \Rightarrow n - x_j = x_1 + \dots + x_{j-1} + x_{j+1} + \dots + x_K$

<u>note</u> $p_1 + \dots + p_K = 1$

$\Rightarrow p_1 + \dots + p_{j-1} + p_{j+1} + \dots + p_K = 1 - p_j$

div both by $1-p_j$ $\Rightarrow \frac{p_1}{1-p_j} + \dots + \frac{p_{j-1}}{1-p_j} + \frac{p_{j+1}}{1-p_j} + \dots + \frac{p_K}{1-p_j} = 1$

$p_1' + \dots p_{j-1}' + p_{j+1}' + \dots + p_K' = 1$

$$= \underbrace{\frac{n'!}{x_1! \dots x_{j-1}!\, x_{j+1}! \dots x_K!} \mathbb{1}_{x_1 + \dots x_{j-1} + x_{j+1} + \dots + x_K = n'} \mathbb{1}_{x_1 \in J_n} \dots \mathbb{1}_{x_{j-1} \in J_n} \dots \mathbb{1}_{x_K \in J_n}}_{\binom{n'}{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_K}} \mathbb{1}^n_{x_j \in J_n} \underbrace{\frac{p_1^{x_1} \dots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \dots p_K^{x_K}}{(1-p_j)^{x_1 + \dots + x_{j-1} + x_{j+1} + \dots + x_K}}}_{p_1'^{x_1} \dots p_{j-1}'^{x_{j-1}} p_{j+1}'^{x_{j+1}} \dots p_K'^{x_K}}$$

$$= \mathrm{Multin}_{K-1}(n', \underbrace{\vec{p}'}_{\begin{bmatrix} p_1' \\ \vdots \\ p_K' \end{bmatrix}}) \mathbb{1}^n_{x_j \in J_n}$$

$\vec{X} \sim \mathrm{Multin}_K(n, \vec{p})$ what is $E[\vec{X}]$? $Var[\vec{X}]$?

241

X is a scalar rv

* $E[aX + c] = aE[X] + c$

* $E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i] = n\mu$ if identically distributed

* $E[\prod_{i=1}^n X_i] = \prod_{i=1}^n E[X_i]$ if they're independent

* $\sigma^2 := Var[X] := E[(X-\mu)^2]$ $\overset{\text{discrete}}{=} \sum_{x \in \mathbb{R}} (x-\mu)^2 p(x)$ $\overset{\text{continuous}}{=} \int_{\mathbb{R}} (x-\mu)^2 f(x) dx$, $\sigma := \sqrt{Var[X]} = SD[X]$ standard deviation

$Var[X] = E[X^2] - \mu^2$

$Var[X_1 + X_2] = E[((X_1 + X_2) - (\mu_1 + \mu_2))^2] = E[X_1^2 + X_2^2 + \mu_1^2 + \mu_2^2 - 2\mu_1 X_1 - 2\mu_1 X_2 - 2\mu_2 X_1 - 2\mu_2 X_2 + 2X_1X_2 + 2\mu_1\mu_2]$

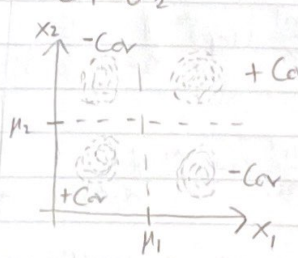
$= E[X_1^2] + E[X_2^2] + \mu_1^2 + \mu_2^2 - 2\mu_1^2 - \cancel{2\mu_1\mu_2} - 2\mu_2\mu_1 - 2\mu_2^2 + 2E[X_1X_2] + \cancel{2\mu_1\mu_2}$

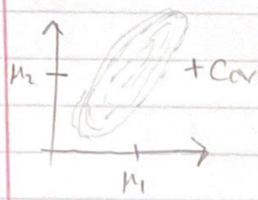
$= \sigma_1^2 + \sigma_2^2 + 2(E[X_1X_2] - \mu_1\mu_2)$ $\sigma_{12} := Cov[X_1, X_2] := E[X_1X_2] - \mu_1\mu_2$

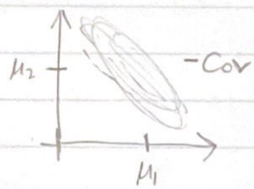
$= \sigma_1^2 + \sigma_2^2 + 2\sigma_{12}$,

if X_1, X_2 indep,
$Cov[X_1, X_2] = \mu_1\mu_2 - \mu_1\mu_2 = 0$

if X_1, X_2 indep,
$= \sigma_1^2 + \sigma_2^2$







Rules for Covariance:

1) $Cov[X,X] = \sigma^2$ (var)
2) $Cov[X_1, X_2] = cov[X_2, X_1]$
3) $Cov[X_1+X_2, X_3] = \sigma_{13} + \sigma_{23}$ (cov)
4) $Cov[a_1X_1, a_2X_2] = a_1a_2\sigma_{12}$ $\quad a_1, a_2 \in \mathbb{R}$

* 5) $Var[X_1 + \ldots + X_n] = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov[X_i, X_j]$

$$\vec{\mu} := E[\vec{X}] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_k] \end{bmatrix}, \text{ let } m := \underbrace{\begin{bmatrix} X_{11} & \cdots & X_{1m} \\ X_{21} & \cdots & X_{2m} \\ \vdots & & \vdots \\ X_{n1} & \cdots & X_{nm} \end{bmatrix}}_{\text{matrix of r.v.'s}}, E[m] = \begin{bmatrix} E[X_{11}] & \cdots & E[X_{1m}] \\ \vdots & & \\ E[X_{n1}] & \cdots & E[X_{nm}] \end{bmatrix}$$

capital sigma

$$\Sigma := Var[\vec{X}] := E[\underbrace{\vec{X}\vec{X}^T}] - \underbrace{\vec{\mu}\vec{\mu}^T}_{\substack{k\times k \\ (k\times1)(1\times k)}} \overset{HW}{=} \begin{bmatrix} Var[X_1] & Cov[X_1,X_2] & \cdots & Cov[X_1,X_k] \\ Cov[X_2,X_1] & Var[X_2] & \cdots & Cov[X_2,X_k] \\ \vdots & & \ddots & \vdots \\ Cov[X_k,X_1] & Cov[X_k,X_2] & \cdots & Var[X_k] \end{bmatrix}$$

outer products

The "variance-covariance matrix". It's square $k \times k$ and symmetric

if $X_1, \ldots, X_k$ are independent then the varcov matrix is:

$$Var[\vec{X}] = diag\{\sigma_1^2, \ldots, \sigma_k^2\} = \begin{bmatrix} \sigma_1^2 & & 0 & 0 \\ 0 & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_k^2 \end{bmatrix}$$

Rules for expectations of vector r.v.'s. Let $\vec{a} \in \mathbb{R}^k$

1) $E[\vec{X} + \vec{a}] = \vec{\mu} + \vec{a}$
2) $E[\vec{a}^T \cdot \vec{X}] = E[a_1X_1 + a_2X_2 + \ldots + a_kX_k] = a_1\mu_1 + \ldots + a_k\mu_k = \vec{a}^T\vec{\mu}$