$\vec{X} \sim Multin_K(n, \vec{p})$

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

$$Deg(n-x_2) = p_{X_1|X_2}(x_1, x_2) := P(X_1 = x_1 \mid X_2 = x_2) = \frac{p(x_1, x_2)}{p(x_2)}$$

$$p(x_2) = Bin(n, p_2) = Bin(n, 1-p_1)$$

$$= \frac{\binom{n}{x_1, x_2} p_1^{x_1} p_2^{x_2}}{\binom{n}{x_2} p_2^{x_2} (\underbrace{1-p_2}_{p_1})^{n-x_2}}$$

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

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

$\mathbb{1}^u_{x_2 \in J_n}$

Define a ratio of indicators:

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

$P(A|B) = \frac{P(A,B)}{P(B)}$ which is undefined if $P(B) = 0$

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

Let's generalize this result a little bit. $\vec{X} \sim Multin_K(n, \vec{p})$, $\vec{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_K \end{bmatrix}$

$$p_{\vec{X}_{-j}|X_j}(\vec{x}_{-j}, x_j) = \frac{p_{\vec{X}}(\vec{x})}{p_{X_j}(x_j)} \sim Multin_{K-1}(n - x_j, ?)$$

All elements of vector rv X except the jth component

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

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

$\mathbb{1}^u_{x_j \in J_n}$

let $n' = n - x_j$. Note $x_1 + \dots + x_K = n \Rightarrow n - x_j = x_1 + \dots + x_{j-1} + x_{j+1} + \dots + x_K$

Note: $p_1 + \dots + p_K = 1$

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

divide both sides by
1 - p_j

$$\Rightarrow \underbrace{\frac{p_1}{1-p_j}}_{p_1'} + \dots + \underbrace{\frac{p_{j-1}}{1-p_j}}_{p_{j-1}'} + \underbrace{\frac{p_{j+1}}{1-p_j}}_{p_{j+1}'} + \dots + \underbrace{\frac{p_K}{1-p_j}}_{p_K'} = 1$$

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

$$= \overbrace{\frac{n'!}{x_1! \cdots x_{j-1}!\, x_{j+1}! \cdots x_K!}}^{\binom{n'}{x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_K}} \mathbb{1}_{x_1 + \dots + x_{j-1} + x_{j+1} + \dots + x_K = n'} \mathbb{1}_{x_1 \in J_n} \cdots \mathbb{1}_{x_{j-1} \in J_n} \cdots \mathbb{1}_{x_K \in J_n}$$

$$\mathbb{1}^u_{x_j \in J_n} \underbrace{\frac{p_1^{x_1} \cdots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \cdots p_K^{x_K}}{(1-p_j)^{x_1 + \dots + x_{j-1} + x_{j+1} + \dots + x_K}}}_{p_1'^{x_1} \cdots p_{j-1}'^{x_{j-1}} p_{j+1}'^{x_{j+1}} \cdots p_K'^{x_K}} = Multin_{K-1}(n', \vec{p}')\, \mathbb{1}^u_{x_j \in J_n}, \quad \vec{p}' = \begin{bmatrix} p_1' \\ \vdots \\ p_K' \end{bmatrix}$$

$\vec{X} \sim Multin_K(n, \vec{p})$ What is $E[\vec{X}]$? $Var[\vec{X}]$?

We need definitions for expectation and variance for vector rv's.

X is a scalar rv. $a, c \in \mathbb{R}$

* $E[aX + c] = aE[X] + c$
* $E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = n\mu$ if they're identically distributed
* $E\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n E[X_i]$ if they're independent

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

standard deviation
$\sigma := \sqrt{Var[X]} = SD[X]$

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

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

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

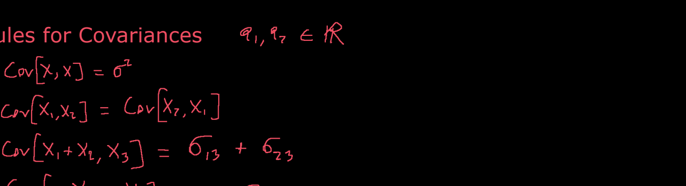
$$= E[X_1^2] + E[X_2^2] + \cancel{\mu_1^2} + \cancel{\mu_2^2} - \cancel{2}\mu_1^2 - \cancel{2\mu_1\mu_2} - 2\mu_2\mu_1 - \cancel{2}\mu_2^2 + 2E[X_1X_2] + \cancel{2\mu_1\mu_2}$$

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

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

If X_1, X_2 independent $\Rightarrow Cov[X_1, X_2] = \mu_1\mu_2 - \mu_1\mu_2 = 0$

$= \sigma_1^2 + \sigma_2^2$ If X_1, X_2 are independent



Rules for Covariances $a_1, a_2 \in \mathbb{R}$

1. $Cov[X, X] = \sigma^2$
2. $Cov[X_1, X_2] = Cov[X_2, X_1]$
3. $Cov[X_1 + X_2, X_3] = \sigma_{13} + \sigma_{23}$
4. $Cov[a_1X_1, a_2X_2] = a_1a_2\sigma_{12}$
5. $Var[X_1 + \dots + X_n] = \sum_{i=1}^n \sum_{j=1}^n Cov[X_i, X_j]$

M is a matrix of rv's

$$\vec{\mu} := E[\vec{X}] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_K] \end{bmatrix}, \quad \text{let } M := \begin{bmatrix} X_{11} & \cdots & X_{1m} \\ X_{21} & \cdots & X_{2m} \\ \vdots & & \\ X_{n1} & \cdots & X_{nm} \end{bmatrix}, \; E[M] = \begin{bmatrix} E[X_{11}] & \cdots & E[X_{1m}] \\ \vdots & & \\ E[X_{n1}] & \cdots & E[X_{nm}] \end{bmatrix}$$

capital sigma

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

outer products

The "variance-covariance matrix". It's square K x K and symmetric.

If X_1, ..., X_K are independent then the varcov matrix is:

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

Rules for expectations of vector rv's. Let $\vec{a} \in \mathbb{R}^K$

1. $E[\vec{X} + \vec{a}] = \vec{\mu} + \vec{a}$
2. $E[\vec{a}^T \vec{X}] = E[a_1X_1 + a_2X_2 + \dots + a_KX_K] = a_1\mu_1 + \dots + a_K\mu_K = \vec{a}^T\vec{\mu}$