$X \sim \text{Multin}_2(n, \vec{p})$

$K=2$

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

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

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

Define $J_n := \{0, 1, \ldots, n\}$

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

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

$$= \underbrace{\frac{(n-x_2)!}{x_1!}}_{\substack{=1 \text{ if} \\ x_1 = n-x_2}} \mathbb{1}_{x_1 = n-x_2} \underbrace{\mathbb{1}_{x_1 \in J_n}}_{\substack{=1 \\ \text{if } x_1 = n-x_2}} \underbrace{p_1^{x_1+x_2-n}}_{\substack{=1 \\ \text{if } x_1 = n-x_2}} \mathbb{1}^u_{x_2 \in J_n} = \text{Deg}(n-x_2)\, \mathbb{1}^u_{x_2 \in J_n}$$

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

Let's generalize this conditional probability a little bit: $\vec{X} \sim \text{Multin}_K(n, \vec{p})$

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

this is the vector without the jth component

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

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

divide both sides by 1-p_j

Note: $p_1 + \ldots + p_K = 1 \Rightarrow p_1 + \ldots + p_{j-1} + p_{j+1} + \ldots + p_K = 1 - p_j$

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

Let $n' := n - x_j$

Note: $n - x_j = x_1 + \ldots + x_{j-1} + x_{j+1} + \ldots + x_K$ o/t prob zero

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

$$\underbrace{\frac{p_1^{x_1} \cdots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \cdots p_K^{x_K}}{(1-p_j)^{x_1+\ldots+x_{j-1}+x_{j+1}+\ldots+x_K}}}_{p_1'^{x_1} \cdots p_{j-1}'^{x_{j-1}} p_{j+1}'^{x_{j+1}} \cdots p_K'^{x_K}} \mathbb{1}^u_{x_j \in J_n}$$

$$= \text{Multin}_{K-1}(n', \vec{p}\,') \mathbb{1}^u_{x_j \in J_n}$$

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

Review from Math 241. Let X_1, ..., X_n be rv's and $a, c \in \mathbb{R}$

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

$$E\left[\sum X_i\right] = \sum E[X_i] \overset{\text{if identically distributed}}{=} n\mu$$

; if independent

$$E\left[\prod X_i\right] = \prod E[X_i]$$

standard deviation

$$\sigma^2 := \text{Var}[X] := E[(X-\mu)^2] = E[X^2] - \mu^2, \quad \sigma := SD[X] := \sqrt{\text{Var}[X]}$$

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

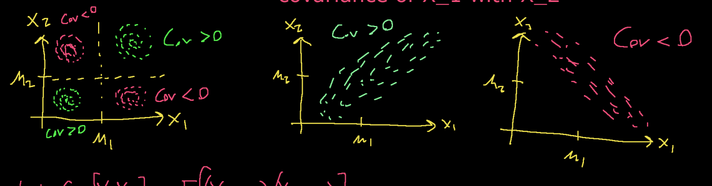
$$= E[X_1^2 + X_2^2 + \mu_1^2 + \mu_2^2 + 2X_1X_2 - 2X_1\mu_1 - 2X_1\mu_2 - 2X_2\mu_1 - 2X_2\mu_2 + 2\mu_1\mu_2]$$

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

$$= \sigma_1^2 + \cancel{\mu_1^2} + \sigma_2^2 + \cancel{\mu_2^2} + \cancel{\mu_1^2} + \cancel{\mu_2^2} + 2E[X_1X_2] - \cancel{2\mu_1^2} - 2\mu_1\mu_2 - \cancel{2\mu_2^2}$$

$$= \sigma_1^2 + \sigma_2^2 + 2(\underbrace{E[X_1X_2] - \mu_1\mu_2}_{\text{Cov}(X_1,X_2)}) = \sigma_1^2 + \sigma_2^2 + 2\underbrace{\text{Cov}[X_1,X_2]}_{\sigma_{12}} \overset{\text{if } X_1, X_2 \text{ independent}}{=} \sigma_1^2 + \sigma_2^2$$

covariance of X_1 with X_2



HW: $\text{Cov}[X_1,X_2] = E[(X_1-\mu_1)(X_2-\mu_2)]$

Covariance Rules:

$$\text{Cov}[X,X] = \sigma^2$$

$$\text{Cov}[X_1,X_2] = \text{Cov}[X_2,X_1]$$

$$\text{Cov}[X_1+X_2, X_3] = \text{Cov}[X_1,X_3] + \text{Cov}[X_2,X_3]$$

$$\text{Cov}[a_1X_1, a_2X_2] = a_1a_2\sigma_{12}$$

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

$$\vec{\mu} := E[\vec{X}] := \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_K] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_K \end{bmatrix}, \quad \text{let } M = \begin{bmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & & \\ X_{n1} & \cdots & X_{nm} \end{bmatrix}$$

$$E[M] := \begin{bmatrix} \mu_{11} & \cdots & \mu_{1m} \\ \vdots & & \\ \mu_{n1} & \cdots & \mu_{nm} \end{bmatrix}$$

Capital letter sigma

$(K\times 1)(1\times K) = K\times K$

$$\Sigma := \text{Var}[\vec{X}] := E[\vec{X}\vec{X}^T] - \vec{\mu}\vec{\mu}^T = \begin{bmatrix} \text{Var}[X_1] & \text{Cov}[X_1,X_2] & \cdots & \text{Cov}[X_1,X_K] \\ \text{Cov}[X_2,X_1] & \text{Var}[X_2] & \cdots & \text{Cov}[X_2,X_K] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[X_K,X_1] & \text{Cov}[X_K,X_2] & \cdots & \text{Var}[X_K] \end{bmatrix}$$

outer product

$K \times K$

variance-covariance (varcov) matrix and it is symmetric

If X_1, ..., X_K are independent, what is the varcov matrix?

$$\Sigma = \text{diag}\{\sigma_1^2, \ldots, \sigma_K^2\} := \begin{bmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots \; \sigma_K^2 \end{bmatrix}$$

Rules about vector rv expectations

$$(\vec{v}_1^T \vec{v}_2)^T = \vec{v}_2^T \vec{v}_1 = \vec{v}_1^T \vec{v}_2$$

$$E[aX + \vec{c}] = \begin{bmatrix} a\mu_1 + c_1 \\ a\mu_2 + c_2 \\ \vdots \\ a\mu_K + c_K \end{bmatrix} = a\vec{\mu} + \vec{c}$$

$$E[\vec{a}^T X] = E[a_1X_1 + \ldots + a_KX_K] = a_1\mu_1 + \ldots + a_K\mu_K = \vec{a}^T\vec{\mu}$$