Bag of fruits: Apples, Banana, and cantaloupe

P_1: Prob of Apples, P_2: Prob of Banana, P_3: Prob of cantaloupe

$P_1 + P_2 + P_3 = 1$

Draw with replacement

Let $X_1 \doteq$ # of Apples, $X_2 \doteq$ # of bananas

$X_3 =$ # of cantaloupes then, $x_1 + x_2 + x_3 = n$

$$\bar{X} = \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} \sim P_{\bar{X}}(\bar{x}) = \frac{n!}{x_1!\, x_2!\, x_3!} P_1^{x_1} P_2^{x_2} P_3^{x_3} \mathbb{1}_{x_1+x_2+x_3=n}$$

$$\mathbb{1}_{x_1 \in \{0,1,\cdots,n\}} \mathbb{1}_{x_2 \in \{0,1,n\}} \mathbb{1}_{x_3 \in \{0,1,\cdots n\}}$$

$$P_{\bar{X}}(\bar{x}) = \binom{n}{x_1\, x_2\, x_3} P_1^{x_1} P_2^{x_2} P_3^{x_3} = \text{multinomial}\left(n, \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}\right)$$

C A C B A B C

Generally with K types of object

$$\bar{X} \sim \text{multinomial}(n, \vec{P}) = \binom{n}{x_1\, x_2\, x_3} P_1^{x_1} P_2^{x_2} P_3^{x_3} \cdots P_K^{x_K}$$

$$\text{supp}[\bar{X}] = \{\vec{x} : \underbrace{\vec{x} \in \mathbb{N}^K}_{\vec{x} \in \{0,1,\cdots,n\}^K}, \vec{x} \cdot \vec{1} = n\} \rightarrow x_1 + x_2 = n$$

$x_1 + 7 = 10$

$$\vec{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \underline{\vec{P} \in \{\vec{P} : \{0,1\}^K, \vec{P} \cdot \vec{1} = 1\}} \quad n \in \mathbb{N}$$

When $n = 2$

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \text{multinomial}\left(n, \begin{bmatrix} P \\ 1-P \end{bmatrix}\right)$$

$P_1 = P$, $P_2 = 1 - P$

$X_1 \sim \text{Binm}(n, P)$

$X_2 \sim \text{Binm}(n, 1-P)$

Is $X_1 \overset{\text{ind}}{=} X_2$? No, Is X_1, X_2 ind? No

For two random variables

$$P(X_1 = x_1 \mid X_2 = x_2) = P(X_1 = x_1) \; \forall \vec{x}_1 \in \text{supp}[X_1] \; \& \; \forall x_2 \in \text{supp}[X_2]$$

$0 = P(X_1 = 1 \mid X_2 = n) \neq P(X_1 = 1) = np(1-p)^{n-1}$

$P_{X_1|X_2}(x_1, x_2) : P(X_1 = x_1 \mid X_2 = x_2) = \frac{P_{X_1,X_2}(x_1, x_2)}{P_{X_2}(x_2)}$

def of conditional probability

$$\frac{\frac{n!}{x_1!\,x_2!} p^{x_1}(1-p)^{x_2}\mathbb{1}_{x_1+x_2=n}}{\frac{n!}{x_2!\,(n-x_2)!} p^{n-x_2}(1-p)^{x_2}} = \frac{(n-x_2)!}{x_1!} p^{x_1+x_2-n}\mathbb{1}_{x_1+x_2=n}$$

$$= \begin{cases} \frac{x_1!}{x_2!} p^0 = 1 & \text{if } x_1+x_2 = n = \text{Deg} \quad (n-x_2) = n - x_2 \text{ w P } 1 \\ 0 & \text{if } x_1 + x_2 \neq n \end{cases}$$

$P_{X_2}(x_2) = \sum_{x_1 \in supp[X_1]} P_{X_1,X_2}(x_1, x_2)$

marginal PMF

marginalization

$P_{X_1,P_2}(x_1\,x_2)$

	x_1: 0	1	2	3	4	5
x_2: 0						
1						
2						

$$= \sum_{x_1 \in \{0,1,\dots,n\}} \frac{n!}{x_1!\,x_2!} p^{x_1}(1-p)^{x_2}\mathbb{1}_{\substack{x_1+x_2=n \\ x_2 \in \{0,1,\dots,n\}}}$$

$$= \sum \frac{n!}{x_1!\,x_2!}\mathbb{1}_{x_1+x_2=n}\mathbb{1}_{x_2\in\{0,1,\dots,n\}}$$

$$= \frac{n!}{x_2!}(1-p)^{x_2}\mathbb{1}_{x_2\in\{0,1,\dots n\}} \sum_{x_1\in\{0,1,\dots n\}} \frac{1}{x_1!} p^{x_1}\mathbb{1}_{x_1 = n-x_2}$$

$= \binom{n}{x_2}(1-p)^{x_2} p^{n-x_2} = \text{Bin}(n, 1-p)$ Thus, the marginal of the multinomial is Binomial.

r.v. $X_1 \mid X_2$

$$\boxed{\text{PMF } p_{X_1|X_2}(x_1,x_2) := P(X_1 = x \mid X_2 = x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)}}$$

$$P_{X_2} = \sum_{x_1\in supp[X_1]} P(x_1,x_2) = \sum_{x_1=0}^{n}\binom{n}{x_1, x_2} p^{x_1}(1-p)^{x_2}\mathbb{1}_{x_1+x_2=0}$$

$P_{\vec{X}_{-j}|X_j}(\vec{x}_{-j}, x_j)$ .... $\vec{X} := \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \cancel{X_k} \end{bmatrix}$

$$= \frac{\frac{n!}{x_1!\,x_{j-1}!\,x_j!\,x_{j+1}!\cdots x_k!} p_1^{x_1}\cdots p_{j-1}^{x_{j-1}} p_j^{x_j} p_{j+1}^{x_{j+1}}\cdots p_k^{x_k}}{\frac{n!}{x_j!\,(n-x_j)!} p_j^{x_j}(1-p_j)^{n-x_j}}$$

Scanned with CamScanner

③

Let $n' = n - x_j$

$$= \frac{n!}{x_1! \cdots x_{j-1} + x_{j+1}\, x_k!} \cdot \frac{p_1^{x_1} \cdots p_{j-1}^{x_{j-1}}\, p_{j+1}^{x_{j+1}} \cdots p_k^{x_k}}{(1-p)^{n'}}$$

$$= (x_1 \cdots x_j, x_{j+1} \cdots x_k) \left(\frac{p_1}{(1-p_j)}\right)^{x_1} \cdots \left(\frac{p_{j-1}}{1-p_j}\right)^{x_{j-1}}$$

$$= \left(\frac{p_{j+1}}{(1-p_j)}\right)^{x_{j+1}} \cancel{x_{j+1}} \cdots \left(\frac{p_k}{(1-p_j)}\right)^{x_k} = \text{multinomial } (n, \vec{p})$$

where $\vec{p}\,' = \begin{bmatrix} \frac{p}{1-p_j} \\ \vdots \\ \frac{p_k}{1-p_j} \end{bmatrix} \rightarrow \dim[\vec{p}] = k-1$ no x_j

$n' = n - x_j$ $\quad n' = x_1 + x_2 + \cdots + x_{j-1} + x_{j+1} + \cdots + x_k$

$n' = x_1 + x_2 + \cdots + x_{j-1} + x_j + x_{j+1} \cdots x_k$

$E[\vec{X}] = ?$ $\quad var[\vec{X}] = ?$ $\quad$ define $\mu = E(X)$

If X is discrete $E(X) = \sum_{x \in R} x\, P(x) \rightarrow$ pmf

If X is continuous $E[X] = \int_R x f(x)\, dx$ (cov)

$E[aX + c] = aE(X) + c = a\mu + c$ where a, c constants

$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i)$ always true

$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E(X_i)$

if $X_1, \ldots X_n \overset{ind}{\sim}$

variance $\sigma^2 = var[X] = E[(X-\mu)^2]$ $\quad \sigma := SE(X) = \sqrt{\sigma^2}$ Standard error

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

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

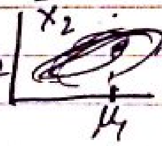
$= E[X_1^2] + E[X_2^2] + \mu_1^2 + \mu_2^2 - 2\mu_1^2 - 2\mu_2\mu_1 - 2\mu_1\mu_2 - 2\mu_1 + 2E[X_1, X_2] + 2\mu_1\mu_2$

$= \sigma_1^2 + \sigma_2^2 + 2[E(X_1, X_2) - \mu_1\mu_2$

$= \sigma_1^2 + \sigma_2^2 + 2[E(X_1, X_2] - \mu_1\mu_2$, if $X_1, X_2 \overset{ind}{\sim} \sigma_1^2 + \sigma_2^2$

$\sigma_{12} = cov[X_1, X_2] := [X_1, X_2] - \mu_1\mu_2 = E[(X_1 - \mu_1)(X_2 - \mu_2)]$

co variance (+)



cov (−)