Bag of Fruits

9/9/2019

p_1 : Prob of Apple

p_2 : Prob of Banana

p_3 : Prob of Cantalope

$p_1 + p_2 + p_3 = 1$

x_k : # of fruits k. A B C A B C C

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim \underset{\text{JMF}}{P_{\vec{X}}(\vec{X})} = \frac{n!}{x_1!\,x_2!\,x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \mathbb{1}_{x_1+x_2+x_3=n} \mathbb{1}_{x_1 \in \{0,1,\dots,n\}}$$

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

$$= \binom{n}{x_1, x_2, x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3} = \text{multinomial}\left(n, \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}\right)$$

Generally with k cathegory

$$\vec{X} \sim \text{multinomial}(n, \vec{p}) = \binom{n}{x_1, x_2, \dots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

$$\vec{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\text{Support}[\vec{X}] = \{\vec{x} : \vec{x} \in \{0, 1, \dots, n\}^k \; \vec{x} \cdot \vec{1} = n\}$$

$$n \in \mathbb{N}, \; \vec{p} \in \{\vec{v} : \vec{v} \in \{0,1\}^k, \; \vec{v} \cdot \vec{1} = 1\}$$

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \text{Multinomial}\left(n, \begin{bmatrix} p \\ 1-p \end{bmatrix}\right) = \binom{n}{x_1, x_2} p^{x_1} (1-p)^{x_2}$$

Bag of Fruit

p_1 : prob of apple

p_2 : prob of banana

$p_1 + p_2 = 1$

$X_1 \sim \text{Bin}(n, p)$

$X_2 \sim \text{Bin}(n, 1-p)$

$X_1 \not\perp X_2$

Are X_1, X_2 ind ?

If so, then $P(X_1 = x_1 \mid X_2 = x_2) = P(X_1 = x_1) \;\; \forall \vec{x} \in \text{Supp}[\vec{X}]$

bananas #draw - #of apples

$$0 = P(X_1 = 1 \mid X_2 = n) \neq P(X_1 = 1) = \underbrace{n p (1-p)^{n-1}}_{\text{PMF Binomial}}$$

$\Rightarrow$ they're dependent

Conditional PMF

tells you X_2, no Randomness in X_1

random quantity X_1

Def. of conditional probability

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

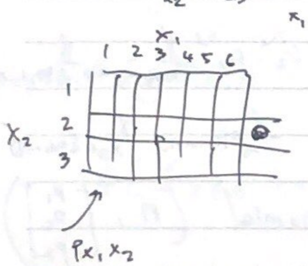
degenerate function

we should know $= \text{Deg}(n - x_2)$ ...it has to be

$= \mathbb{1}_{x_1 = n - x_2}$

Marginal PMF

$$p_{X_2}(x_2) = \sum_{x_1 \in \text{Supp}[X_1]} p_{X_1 X_2}(x_1, x_2) = \sum_{x_1 \in \{0, 1, \dots, n\}} \frac{n!}{x_1! \, x_2!} p^{x_1} (1-p)^{x_2} \mathbb{1}_{x_1 + x_2 = n}$$



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

i.e. $\sum_{i=1}^{100} i \, \mathbb{1}_{i=17} = 17$

$$= \frac{n!}{x_2!}(1-p)^{x_2} \frac{1}{(n-x_2)!} p^{n-x_2}$$

$$= \binom{n}{x_2}(1-p)^{x_2} p^{n-x_2}$$

$$= \text{Bin}(n, 1-p)$$

$\vec{x} \sim \text{Bin}\ \cancel{\text{Multin}}(n, \vec{p})$

$\Rightarrow x_j \sim \text{Bin}(n, p_j)$

$\text{Var}[\vec{x}] = np_1(1-p_1) \ldots$

---

multinomial PMF

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

binomial PMF

$$= \frac{(n-x_2)!}{x_1!} p^{x_1+x_2-n} \mathbb{1}_{x_1+x_2=n} = \begin{cases} \frac{x_1!}{x_1!} p^0 = 1 & \text{if } x_1 = n - x_2 \Leftrightarrow x_1 + x_2 = n \\ 0 & \text{o/t} \end{cases}$$

$$= \text{Deg}(n - x_2) = \{ n - x_2 \text{ w.p. } 1$$

$$P_{\vec{x}_{-j} \mid x_j}(\vec{x}_{-j}, x_j) = \frac{\text{Multinomial}(n, \vec{p})}{\text{Bin}(n, p_j)}$$

Let $\vec{x}_{-j} = \begin{bmatrix} x_1 \\ \vdots \\ x_{j-1} \\ x_{j+1} \\ \vdots \\ x_k \end{bmatrix} \Rightarrow \dfrac{\dfrac{n!}{x_1! \ldots x_{j-1}!\,x_j!\,x_{j+1}! \ldots x_k!} p_1^{x_1} \ldots p_{j-1}^{x_{j-1}} p_j^{x_j} p_{j+1}^{x_{j+1}} \ldots p_k^{x_k}}{\dfrac{n!}{x_j!\,(n-x_j)!} p_j^{x_j} (1-p_j)^{n-x_j}}$

Let $n' = n - x_j \Rightarrow \dfrac{n'!}{x_1! \ldots x_{j-1}!\,x_{j+1}! \ldots x_k!} \dfrac{p_1^{x_1} \ldots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \ldots p_k^{x_k}}{(1-p_j)^{n'}}$

$(1-p_j)^{n'} = (1-p_j)^{x_1} \ldots (1-p_j)^{x_{j-1}} (1-p_j)^{x_{j+1}} \ldots (1-p_j)^{x_k}$

note $n = x_1 + \ldots + x_{j-1} + x_j + x_{j+1} + \ldots + x_k$

$n' = x_1 + \ldots + x_{j-1} + x_{j+1} + \ldots + x_k$

$\Rightarrow \text{Multinomial}(n', \vec{p}\,')$ when

$\vec{p} = \begin{bmatrix} \frac{p_1}{1-p_j} \\ \vdots \\ \frac{p_{j-1}}{1-p_j} \\ \frac{p_{j+1}}{1-p_j} \\ \vdots \end{bmatrix}$ such that $\dim[\vec{p}] = k-1$

$\frac{p_k}{1-p_j} \rightarrow$

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix} \quad E[\vec{X}] = ? \quad Var[\vec{X}] = ?$$

Let $X_1, \ldots, X_n$ be r.v's

$E[aX + c] = a\mu + c$ where $\mu = E[X]$, $a, c \in \mathbb{R}$ constant

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

$\overset{discrete}{=} \sum_{x \in \mathbb{R}} (x-\mu)^2 p(x)$

$\overset{constant}{=} \int_{\mathbb{R}} (x-\mu)^2 f(x)\,dx$

$\sigma^2 = Var[X] = E[(X-\mu)^2] = \ldots = E[X^2] - \mu^2$

$\sigma = SE[X] = \sqrt{\sigma^2}$

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

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

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

$\sigma_1^2 \quad \sigma_2^2$

$\Rightarrow \sigma_1^2 + \sigma_2^2 + 2(E[X_1X_2] - \mu_1\mu_2)$

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

Covariance

$\sigma_{12} = Cov[X_1, X_2] = E[X_1X_2] - \mu_1\mu_2 = E[(X_1 - \mu_1)(X_2 - \mu_2)]$