Bag of Fruit: Apples and Bannanas and Cantaloupe only

p_1: prob. of Apple

p_2: prob. of Bannana $\qquad p_1 + p_2 + p_3 = 1$

p_3: prob of Cantaloupe

Draw n with replacement

Let X_1 = # of apples

X_2 = # of bannanas $\qquad X_1 + X_2 + X_3 = n$

X_3 = # of canteloupe

JMF joint mass function (gives joint probability)

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim \underset{\text{has JMF}}{p_{\vec{X}}(\vec{x})} = \frac{n!}{x_1!\,x_2!\,x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \mathbb{1}$$

$x_1 \in \{0, 1, \dots, n\}$

$x_2 \in \{0, 1, \dots, n\}$

$x_3 \in \{0, 1, \dots, n\}$

$x_1 + x_2 + x_3 = n$

combine

$$p_{\vec{X}}(\vec{x}) = \binom{n}{x_1, x_2, x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

$$= \text{Multinomial}\left(n, \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}\right)$$

Generally with K types of objects

$$\vec{X} \sim \text{Multinomial}(n, \vec{p}) =$$

$$\binom{n}{x_1, x_2, \dots, x_n} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

$$\text{Supp}[\vec{X}] = \{\vec{x} : \vec{x} \in \mathbb{N}^K, \vec{x} \cdot \mathbb{1} = n\}$$

note that here, $\mathbb{N}$ includes 0

Generally with K types of objects

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

$= \binom{n}{x_1, x_2, \ldots, x_n} p_1^{x_1} p_2^{x_2} \cdots p_n^{x_n}$

here, $\mathbb{N}$ includes 0

$\text{Supp}[\vec{X}] = \{\vec{x} : \vec{x} \in \mathbb{N}^K, \vec{x} \cdot \vec{1} = n\}$

$= \{\vec{x} : \vec{x} \in \{0, 1, 2, \ldots, n\}^K, \vec{x} \cdot \vec{1} = n\}$

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

parameter space:

$\vec{p} \in \{\vec{p} : \{0, 1\}^K, \vec{p} \cdot \vec{1} = 1\}$

Case where $n = 2$:

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

$p_1 = p \qquad X_1 \sim \text{Bin}(n, p)$
$p_2 = 1-p \qquad X_2 \sim \text{Bin}(n, 1-p)$

Is $X_1 = X_2$? No

Are X_1 and X_2 identically distributed?

No, have different parameter (p vs. $1-p$)

Are X_1, X_2 independent?

No, if get X_1, know what X_2 is exactly (not unknown).

Proof:

For two indep. r.v.'s

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

r.v.
random variables

Proof:

For 2 indep r.v.'s

$$P(X_1 = x_1 \mid X_2 = x_2) = P(X_1 = x_1) \ \forall \vec{x} \in Supp[\vec{X}]$$

If this is false, X_1 and X_2 are not indep.

indep.
independent

$$\underbrace{P(X_1 = 1 \mid X_2 = n)}_{=0} \stackrel{?}{=} P(X_1 = 1) = np(1-p)^{n-1}$$

$$0 \neq np(1-p)^{n-1}$$

so X_1, X_2 are not indep. (are dependent)

Conditional PMF/JMF

$$\boxed{P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}}$$

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

by Def of conditional probability

marginal PMF $p_{X_1X_2}(x_1, x_2)$

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

this summing to get $p_{X_2}(x_2)$ is called marginalization

in our case

$$p_{X_2}(x_2) = \sum \frac{n!}{x_1!\,x_2!} p^{x_1}(1-p)^{x_2} \mathbb{1}_{x_1+x_2=n}$$

$x_2 \in \{0, 1, \ldots, n\}$

$$p_{X_2}(x_2) = \sum_{x \in Supp[X_1]} p_{X_1,X_2}(x_1, x_2)$$

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

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

$$\underbrace{\sum \frac{1}{x_1!} \mathbb{1}_{x_1 = n - x_2}}$$

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

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

$$= Bin(n, 1-p)$$

We've proved that the marginal of the multinomial is the binomial

going back...

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

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

cancel stuff to get

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

$$= \begin{cases} \frac{x_1!}{x_1!} p^0 = 1 & \text{if } x_1 + x_2 = n \\ 0 & \text{if } x_1 + x_2 = n \end{cases}$$

$$= Deg(n - x_n) = \{n - x_2 \text{ w.p. } 1,\ 0 \text{ o.w.}\}$$

for $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

w.p. means with probability

$$p_{x_1|x_2}(x_1, x_2) = \mathrm{Deg}(n - x_2) = \overset{\text{w.p.}}{\begin{cases} 1 & \text{if } x_1 = n - x_2 \\ 0 & \text{otherwise} \end{cases}}$$

In general, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix}$

$\vec{x}_{-j}$ means all of $\{x_1, x_2, \ldots x_K\}$ except x_j

$$\underset{\substack{\uparrow \\ \text{joint marginal prob. (without } \vec{x}_j)}}{p_{\vec{x}_{-j}|\vec{x}_j}}(\vec{x}_{-j}, x_j) = \frac{\dfrac{n!}{x_1!x_2!\ldots \underset{\text{skip } x_j}{x_{j-1}!x_{j+1}!}\ldots x_K!} p_1^{x_1} p_2^{x_2} \cdots p_{j-1}^{x_{j-1}} \underset{\text{no } p_j}{} p_{j+1}^{x_{j+1}} \cdots p_K^{x_K}}{\dfrac{n!}{x_j!(n-x_j)!} p_j^{x_j}(1-p_j)^{n-x}}$$

$$= \frac{n!}{x_1!x_2!\ldots x_{j-1}!x_{j+1}!\ldots x_K!} \; \frac{p_1^{x_1} p_2^{x_2} \cdots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \cdots p_K^{x_K}}{(1-p_j)^n}$$

$$= \binom{n}{x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots x_K} \left(\frac{p_1}{1-p_j}\right)^{x_1} \left(\frac{p_2}{1-p_j}\right)^{x_2} \cdots \left(\frac{p_{j-1}}{1-p_j}\right)^{x_{j-1}} \underset{\substack{\uparrow \\ \text{no } p_j}}{\left(\frac{p_{j+1}}{1-p_j}\right)^{x_{j+1}}} \cdots \left(\frac{p_K}{1-p_j}\right)^{x_K}$$

$$= \mathrm{Multinomial}(n', \vec{p}\,')$$

where $\vec{p}\,' = \begin{bmatrix} \frac{p_1}{1-p_j} \\ \frac{p_2}{1-p_j} \\ \vdots \\ \frac{p_{j-1}}{1-p_j} \\ \frac{p_{j+1}}{1-p_j} \\ \vdots \\ \frac{p_K}{1-p_j} \end{bmatrix}$ $\quad n' = n - x_j$

$\dim[\vec{p}\,'] = k - 1$

skip p_j (only one not there is for p_j)

$$n = x_1 + x_2 + \ldots + x_{j-1} + x_j + x_{j+1} + \ldots + x_K$$
$$n' = x_1 + x_2 + \ldots + x_{j-1} \underset{\substack{\uparrow \\ \text{no } x_j}}{} + x_{j+1} + \ldots + x_K$$

so $n = n' + x_j$

$n' = n - x_j$

$E[\vec{X}] = ?$ $\quad\quad$ $Var[\vec{X}] = ?$

define $\mu = E(X)$

if X is discrete $E(X) = \sum_{x \in \mathbb{R}} x\,p(x)$ ← $p(x)$ is PMF

if X is continuous $E(X) = \int_{\mathbb{R}} x\,f(x)\,dx$ ← $f(x)$ is PDF

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

$E[X+Y] = E[X] + E[Y]$

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

if X, Y are indep. $E[XY] = E[X] \cdot E[Y]$

if $X_1, X_2, \ldots, X_n$ are independent:

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

if X, Y are indep. $Var[X+Y] = Var[X] + Var[Y]$
Proved on next page

Variance:

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

Standard Error (Standard Deviation)

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

Covariance:

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

$$\mu_{x_1+x_2} = \mu_{x_1} + \mu_{x_2}$$

If have 2 r.v.'s

$$Var[X_1 + X_2] = E[((x_1 + X_2) - (\mu_1 + \mu_2))^2]$$

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

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

Since $E[X_1^2] - \mu_1^2 = \sigma_1^2$

$E[X_2^2] - \mu_2^2 = \sigma_2^2$

we have

$$Var[X_1 + X_2] = \sigma_1^2 + \sigma_2^2 + 2(E[X_1X_2] - \mu_1\mu_2)$$

$$Var[X_1 + X_2] = \sigma_1^2 + \sigma_2^2 + 2\,Cov[X_1, X_2]$$

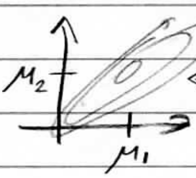
If X_1 and X_2 are independent

$$Var[X_1 + X_2] = \sigma_1^2 + \sigma_2^2$$

$$\sigma_{12} \overset{\text{is defined}}{:=} \underset{\uparrow}{\mathrm{Cov}}[X_1, X_2] := E[X_1 X_2] - \mu_1 \mu_2$$

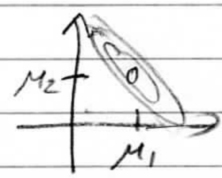
"<u>Covariance</u>"

$$= E\left[(X_1-\mu_1)(X_2-\mu_2)\right]$$



← covariance is (+)

as X_1 higher more likely that X_2 is higher



← covariance is (−)

as X_1 is higher, more likely that X_2 is lower