## Example of a multivariate discrete $\text{dist}^n$

Consider a random experiment with 3 mutually exclusive and exhaustive outcomes, $A_1, A_2$ & $A_3$ with probabilities $\theta_1, \theta_2, \theta_3$, respectively. Repeat the trials $n$ times

Define

$X_1$ : number of times $A_1$ occurs out of $n$ trials

$X_2$ : - - - - - $A_2$ occurs out of $n$ trials

$X_3$ : - - - - - $A_3$ occurs out of $n$ trials let $(x_1, x_2, x_3)$ denote the observed count in n trials

$$x_i \geq 0, \quad x_i \leq n \quad \forall i = 1,2,3 \quad \& \quad \sum_1^3 x_i = n$$

$$E = \left\{ (x_1, x_2, x_3) : 0 \leq x_i \leq n, \ \sum_{i=1}^3 x_i = n \right\}$$ - finite number of points

jt p.m.f.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{n!}{x_1!\, x_2!\, x_3!} \theta_1^{x_1} \theta_2^{x_2} \theta_3^{x_3}; \quad \begin{array}{l} 0 \leq x_i \leq n \\ \sum x_i = n \end{array}$$

Note that $x_3 = n - x_1 - x_2$ and

$$\theta_3 = 1 - \theta_1 - \theta_2$$

$$P(X_1 = x_1, X_2 = x_2) = \frac{n!}{x_1!\, x_2!\, (n - x_1 - x_2)!} \theta_1^{x_1} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{n - x_1 - x_2}$$

$$x_1, x_2 \geq 0$$

$$x_1 + x_2 \leq n$$

$$= 0, \quad \text{o/w}$$

$(X_1, X_2)$ is said to follow a trinomial distⁿ $(n, \theta_1, \theta_2)$

<u>Marginal distⁿs</u>:

Marginal p.m.f. of $X_1$:

$$P(X_1 = x_1) = \sum_{x_2 = 0}^{n - x_1} \frac{n!}{x_1!\,(n - x_1)!} \theta_1^{x_1} \frac{(n - x_1)!}{x_2!\,(n - x_1 - x_2)!} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{n - x_1 - x_2}$$

$$= \binom{n}{x_1} \theta_1^{x_1} (1 - \theta_1 - \theta_2 + \theta_2)^{n - x_1}$$

$$= \binom{n}{x_1} \theta_1^{x_1} (1 - \theta_1)^{n - x_1}.$$

i.e. $X_1 \sim Bin(n, \theta_1)$

Sly $X_2 \sim Bin(n, \theta_2)$

Conditional dist$^n$ of $X_1 | X_2$ :

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

$$= \frac{\frac{\cancel{n!}}{x_1!\, \cancel{x_2!}\, (n - x_1 - x_2)!} \theta_1^{x_1} \cancel{\theta_2^{x_2}} (1 - \theta_1 - \theta_2)^{n - x_1 - x_2}}{\left(\frac{\cancel{n!}}{\cancel{x_2!}\, (n - x_2)!}\right) \cancel{\theta_2^{x_2}} (1 - \theta_2)^{n - x_2}}$$

$$= \frac{(n - x_2)!}{x_1!\, (n - x_2 - x_1)!} \left(\frac{\theta_1}{1 - \theta_2}\right)^{x_1} \left(1 - \frac{\theta_1}{1 - \theta_2}\right)^{n - x_2 - x_1}$$

$x_1 = 0, 1, \ldots, n - x_2$

i.e. $X_1 | X_2 \sim \text{Bin}\left(n - x_2, \frac{\theta_1}{1 - \theta_2}\right)$

Sly $X_2 | X_1 \sim \text{Bin}\left(n - x_1, \frac{\theta_2}{1 - \theta_1}\right)$

Note: Extension to $p > 3$ case – multinomial dist$^n$

$p > 3$ outcomes $A_1, \ldots, A_p$ – mutually exclusive & exhaustive

w.p. $\theta_1, \ldots, \theta_p$ $\quad \sum_1^p \theta_i = 1$ ; $\theta_i \geq 0$

$n$ repeated trials

$X_1$ : number of times $A_1$ occurs

$X_2$ : - - - - - $A_2$ - - -

⋮

$X_p$ : - - - - - $A_p$ - - -

Particular realisation $(x_1, \ldots x_p)$ ; $0 \leq x_i \leq n \quad i = 1, \ldots p$

$\sum_{i=1}^p x_i = n$

$E = \{(x_1, \ldots, x_p) : 0 \leq x_i \leq n, \sum_1^p x_i = n\}$ ← finite number of points

possible values of random vector $\underset{\sim}{X} = (X_1, \ldots, X_p)'$

jt p.m.f.

$$P(X_1=x_1,\ldots,X_p=x_p)=\begin{cases}\dfrac{n!}{x_1!\cdots x_p!}\,\theta_1^{x_1}\theta_2^{x_2}\cdots\theta_p^{x_p}, & x_i\geq 0,\ \sum_1^p x_i=n\\[2ex] 0, & o/w\end{cases}$$

Note that $x_p = n-x_1-\cdots-x_{p-1}$

$\theta_p = 1-\theta_1-\cdots-\theta_{p-1}$

jt p.m.f. of $X_1,\ldots X_{p-1}$

$$P(X_1=x_1,\ldots X_p=x_p)=\begin{cases}\dfrac{n!}{x_1!\,x_2!\cdots(n-x_1-\cdots-x_{p-1})!}\,\theta_1^{x_1}\theta_2^{x_2}\cdots(1-\theta_1-\cdots-\theta_{p-1})^{n-x_1-\cdots-x_{p-1}}, & x_i\geq 0;\ \sum_1^{p-1}x_i\leq n\\[2ex] 0, & o/w.\end{cases}$$

Marginal distr's:

$X_i \sim B(n,\theta_i)\quad i=1,\ldots,p$

jt marginal $(X_i,X_j)\sim \text{trinomial}(n,\theta_i,\theta_j)$

jt marginal $(X_i,X_j,X_k)\sim \text{multinomial}(n,\theta_i,\theta_j,\theta_k)$

Conditional distr's:

$X_i \mid X_j=x_j \sim Bin\left(n-x_j,\ \dfrac{\theta_i}{1-\theta_j}\right)$

$(X_i,X_j)\mid X_k=x_k \sim \text{trinomial}\left(\cancel{n}\,x_k,\ \dfrac{\theta_i}{1-\theta_k},\ \dfrac{\theta_j}{1-\theta_k}\right)$

$\left(n-x_k,\ \dfrac{\theta_i}{1-\theta_k},\ \dfrac{\theta_j}{1-\theta_k}\right)$

Continuous multivariate distributions

A $p$-dimensional random vector $\underset{\sim}{X} = (X_1, \ldots, X_p)'$ is said to be (absolutely) continuous if $\exists$ a $f^n$ $f_{X_1,\ldots X_p}(x_1,\ldots,x_p) \geq 0$ $\ni$ the jt d.f. of $(X_1,\ldots X_p)$ is expressed as

$$F_{X_1,\ldots X_p}(x_1,\ldots x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f_{X_1,\ldots X_p}(t_1,\ldots,t_p)\, dt_1 \ldots dt_p$$

$f_{X_1,\ldots X_p}(x_1,\ldots,x_p)$ is called the jt p.d.f. of $(X_1,\ldots X_p)$

$$f_{X_1,\ldots X_p}(x_1,\ldots x_p) = \frac{\partial^p F_{X_1,\ldots X_p}(x_1,\ldots x_p)}{\partial x_1 \partial x_2 \ldots \partial x_p}$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1\ldots X_p}(x_1,\ldots x_p)\, dx_1 \ldots dx_p = 1$$

Note that all marginal p.d.f.s can be obtained from the joint p.d.f. (or jt d.f.)

## Marginal distⁿs

Marginal ~~dist~~ p.d.f. of $X_1$:

$$f_{X_1}(x_1) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-1 \text{ fold};\ x_2 \to x_p} f_{\underset{\sim}{X}}(x_1,\ldots,x_p)\, dx_2 \ldots dx_p$$

marginal p.d.f. of any $X_i$

$$f_{X_i}(x_i) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-1 \text{ fold; all except } x_i} f_{\underset{\sim}{X}}(x_1,\ldots,x_p)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_p.$$

jt marginal p.d.f. of $X_{i_1}, \dots, X_{i_q}$

$$f_{X_{i_1}, \dots, X_{i_q}}(x_{i_1}, \dots, x_{i_q}) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{p-q \text{ fold; all except } x_{i_1}, \dots, x_{i_q}} f_{\underset{\sim}{X}}(x_1, \dots, x_p)\, d(\ )$$

Conditional dist$^n$s

Conditional p.d.f. of $X_i$ given $X_j$:

$$f_{X_i | X_j}(x_i) = \frac{f_{X_i, X_j}(x_i, x_j)}{f_{X_j}(x_j)} \quad (\text{for } f_{X_j}(x_j) > 0)$$

joint conditional p.d.f. of $(X_1, \dots X_q)$ given $(X_{q+1}, \dots, X_p)$

$$f_{X_1, \dots, X_q | X_{q+1}, \dots, X_p}(x_1, \dots x_q) = \frac{f_{\underset{\sim}{X}}(\underset{\sim}{x})}{f_{X_{q+1}, \dots, X_p}(x_{q+1}, \dots, x_p)}$$

$(\text{for } f_{X_{q+1}, \dots X_p}(x_{q+1}, \dots, x_p) > 0)$

Sly conditional dist$^n$ for any subset given any other subset can be obtained.

Independence

Def$^n$: $(X_1, \dots, X_p)$ are pairwise indep iff

$$\forall i \neq j \qquad f_{X_i, X_j}(x_i, x_j) = f_{X_i}(x_i)\, f_{X_j}(x_j) \quad \forall (x_i, x_j)$$

Def$^n$: $(X_1, \dots, X_p)$ are indep iff

$$f_{X_1, \dots X_p}(x_1, \dots, x_p) = \prod_{i=1}^{p} f_{X_i}(x_i) \qquad \forall\ \underset{\sim}{x}$$

Note: Independence $\Rightarrow$ pair wise indep.

Converse is not true