Multinomial Distribution.

Random Expt. with K possible outcomes at each stage was replicated n times

p_j : prob. of j'th outcome at each stage

$(N_1, N_2, \ldots, N_K)$: vector of frequencies of $\{1, 2, \ldots, k\}$.

Joint pmf. of $(N_1, \ldots, N_K)$:

$$p(N_1 = n_1, N_2 = n_2, \ldots, N_K = n_K)$$

$$= \frac{n!}{\prod_{j=1}^{K} n_j!} \prod_{j=1}^{K} p_j^{n_j}$$

if $n_1 + n_2 + \cdot + n_k = n$

and 0 otherwise

$(N_1, \dots, N_K) \sim \text{Mult}(n, p_1, p_2, \dots, p_K)$

$$(N_1, \dots, N_K) = \left(\sum_{i=1}^{n} N_{i,1}, \dots, \sum_{i=1}^{n} N_{iK}\right)$$

where $(N_{i,1}, \dots, N_{iK})$ is the vector of counts of outcomes in the i'th trial.

---

i'th trial:

$$\boxed{(N_{i1}, \dots, N_{iK})} = (0, 0, \dots, 1, 0, \dots 0)$$

There are K possible values, $e_1, e_2, \dots, e_K$ the K canonical vectors

$$(N_{i1}, \dots, N_{iK}) \sim \text{Mult}(1, p_1, \dots, p_K)$$

$$\begin{pmatrix} N_1 \\ \vdots \\ N_K \end{pmatrix} = \sum_{i=1}^{n} \underbrace{\begin{pmatrix} N_{i,1} \\ \vdots \\ N_{i,K} \end{pmatrix}}$$

indep. and identically distributed

Note that $(N_{i1}, \dots, N_{iK})$ are i.i.d

Fact: Suppose:

$(N_1, \ldots, N_K) \sim \text{Mult}(n, p_1, p_2, \ldots, p_K)$

independent

$(\tilde{N}_1, \tilde{N}_2, \ldots, \tilde{N}_K) \sim \text{Mult}(\tilde{n}, p_1, \ldots, p_K)$

What is the distribution of the sum?

$(N_1 + \tilde{N}_1, N_2 + \tilde{N}_2, \ldots, N_K + \tilde{N}_K)$

$\sim \text{Mult}(n + \tilde{n}, p_1, \ldots, p_K)$

Write out the argument.

Multinomial $(n, p_1, \ldots, p_K)$

Question: What is the distribution of $(N_1, N_2, \ldots, N_M)$ where $M < K$?

$(N_1, N_2, \ldots, N_M, \tilde{N}_{M+1})$

$\tilde{N}_{M+1} = n - (N_1 + N_2 + \cdots + N_M)$

Equivalently interested in:

$(N_1, N_2, \ldots, N_M, \tilde{N}_{M+1} \equiv n - (N_1 + \cdots + N_M))$

Here we're looking at a multinomial with parameters $(n, \underline{p_1, p_2, \ldots, p_M}, \underline{\tilde{p}_{M+1}})$

where $\tilde{p}_{M+1} = 1 - p_1 - p_2 \cdots - p_M$.

$N_1 \sim Bin(n, p_1)$

Conditional distribution of $(N_{M+1}, \ldots, N_K)$ given $\underline{(N_1, \ldots, N_M)} \longrightarrow$ equivalently $(N_1, N_2, \ldots, \tilde{N}_{M+1})$.

Need to find the conditional mass function.

$$P(N_{m+1} = n_{m+1}, N_{m+2} = n_{m+2}, \ldots, N_K = n_K \mid N_1 = n_1, \ldots, N_m = n_m)$$

$$\tilde{N}_{m+1} = n - \left(\sum_{j=1}^{m} n_j\right)$$

$$= 0 \quad \text{if } n_{m+1} + \cdots + n_K \neq n - \sum_{j=1}^{m} n_j$$

Else:

Compute the ratio. The denominator is a multinomial prob.

Check that numerator is

$$P(N_1 = n_1, N_2 = n_2, \ldots, N_K = n_K)$$

which is also multinomial

Conclude that conditional is

$$\text{Mult}\left(n - \sum_{j=1}^{m} n_j, \frac{p_{m+1}}{1 - p_1 - \cdots - p_m}, \ldots, \frac{p_K}{1 - p_1 - \cdots - p_m}\right)$$

Continuous multidimensional random vectors.

A random vector $(X_1, \ldots, X_d)$ is said to be continuous if $F_{X_1, \ldots, X_d}$ is continuous on $\mathbb{R}^d$

$$F_{X_1, \ldots, X_d}(x_1, \ldots, x_d) = P(X_1 \le x_1, \ldots, X_d \le x_d)$$

We'll confine ourselves to continuous random vectors that are called in the ABSOLUTELY CONTINUOUS sense that there is $f_{X_1, X_2, \ldots, X_d}(x_1, \ldots, x_d) = f(x_1, \ldots, x_d) \geq 0$ and such that : probability density function

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_d \in A_d) = \int_{A_1 \times \ldots \times A_d} f(x_1, \ldots, x_d)\, dx_1 \ldots dx_d$$

and A_i's are Borel sets.

Equivalently:

$$P\left((X_1, \dots, X_d) \in B\right) = \int_B f(x_1, x_2, \dots, x_d)\, dx_1 \dots dx_d$$

Equivalently:

$$F_{X_1, \dots, X_d}(x_1, \dots, x_d) = \int_{[-\infty, x_1] \times \dots \times [-\infty, x_d]} f(u_1, \dots, u_d)\, du_1 \dots du_d$$

Restrict to $d = 2$ for the moment.

$$F_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \underline{f(u_1, u_2)}\, du_2\, du_1$$

What about the marginals?

$$\underline{F_{X_1}(x_1)} = \underline{P(X_1 \le x_1, X_2 < \infty)}$$

$$\shortparallel$$

$$P(X_1 \le x_1) = \int_{-\infty}^{x_1} \left[\int_{-\infty}^{\infty} \underline{f(u_1, u_2)}\, du_2 \right] du_1$$

$$= f_{X_1}(u_1)$$

Marginal densities:

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1$$

Conditional probability density functions

I want to talk about the conditional density / conditional distribution of X_2 given $X_1 = x_1$

$P(X_2 \le x_2 \mid X_1 = x_1)$ What's this?

Can't use BAYES RULE.

since $P(X_1 = x_1) = 0$

Need to extend the notion of conditioning.

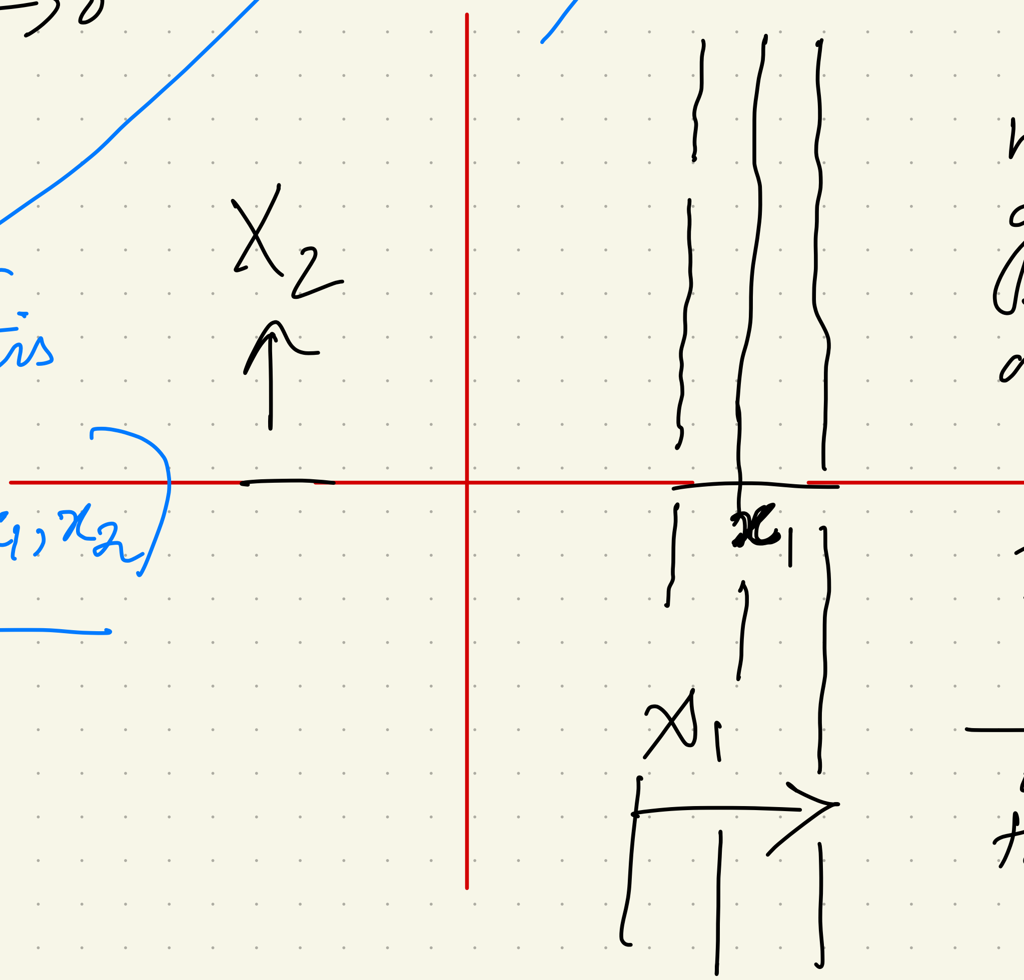
$$P(X_2 \le x_2 \mid X_1 = x_1)$$

We can think of it as:

$$\lim_{h \to 0} P(X_2 \le x_2 \mid X_1 \in (x_1 - h, x_1 + h))$$

Call this

$\underline{P(h, x_1, x_2)}$



We're going to assume that $f_{X_1}(x) > 0$ on the whole $\underline{\text{neighborhood}}$

So:

$$P(X_2 \le x_2 \mid \underline{X_1 \in (x_1 - h, x_1 + h)})$$

$$= \frac{P(X_1 \in (x_1 - h, x_1 + h), X_2 \le x_2)}{P(X_1 \in (x_1 - h, x_1 + h))}$$

$$= \frac{\displaystyle\int_{x_1-h}^{x_1+h} \int_{-\infty}^{x_2} f(u,v)\, dv\, du}{\displaystyle\int_{x_1-h}^{x_1+h} f_{X_1}(u)\, du}$$

Set $G(u, x_2) = \displaystyle\int_{-\infty}^{x_2} f(u,v)\, dv$

So:

$$P(h, x_1, x_2) = \frac{\frac{1}{2h}\displaystyle\int_{x_1-h}^{x_1+h} G(u, x_2)\, du}{\frac{1}{2h}\displaystyle\int_{x_1-h}^{x_1+h} f_{X_1}(u)\, du}$$

$q(u)$

$$\lim_{h\to 0} P(h, x_1, x_2) = \frac{\displaystyle\lim_{h\to 0} \frac{1}{2h}\int_{x_1-h}^{x_1+h} G(u, x_2)\, du}{\displaystyle\lim_{h\to 0} \frac{1}{2h}\int_{x_1-h}^{x_1+h} f_{X_1}(u)}$$

If q is a continuous function from $\mathbb{R} \to \mathbb{R}$, (*)

$$\lim_{h \to 0} \frac{\int_{x-h}^{x+h} q(u)\,du}{2h} = q(x).$$

So assuming continuity,

$$\lim_{h \to 0} \mathbb{P}(h, x_1, x_2)$$

$$= \frac{G(x_1, x_2)}{f_{X_1}(x_1)} \quad \text{(using the result (*))}$$

$$= \int_{-\infty}^{x_2} \frac{f(x_1, v)}{f_{X_1}(x_1)}\,dv$$

$$\int_{-\infty}^{\infty} \psi(v)\,dv = 1$$

$$\underbrace{\mathbb{P}(X_2 \le x_2 \mid X_1 = x_1)}_{\text{conditional dist of } X_2 \text{ given } X_1 = x_1} = \int_{-\infty}^{x_2} \underbrace{\frac{f(x_1, v)}{f_{X_1}(x_1)}}_{\psi(v)}\,dv$$

Conditional density of $X_2 | X_1 = x_1$

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

Similar formula to what we see for probability mass functions.

Independence of the random variables X_1 and X_2: (characterize in terms of densities)

Remember:

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

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n)$$

$$= P(X_1 \in A_1) \cdots P(X_n \in A_n)$$

for all Borel sets $A_1, A_2, \ldots, A_n$.

So in this particular case:

$$F_{X_1, X_2}(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$$

$$= P(X_1 \leq x_1) P(X_2 \leq x_2)$$

if X_1 and X_2 are independent

In which case:

$$F_{X_1, X_2}(x_1, x_2) = \left(\int_{-\infty}^{x_1} f_{X_1}(u)\,du\right)\left(\int_{-\infty}^{x_2} f_{X_2}(v)\,dv\right)$$

for all (x_1, x_2)

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \left(f_{X_1}(u) f_{X_2}(v)\right) dv\,du$$

On the other hand:

$$F_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_{X_1, X_2}(u, v)\,dv\,du$$

This implies ~~that~~: $f_{X_1}(u) f_{X_2}(v) = f_{X_1, X_2}(u, v)$ outside of a negligible set