§29. Many Random Variables; Multinomial

Let $X_1, X_2, \ldots, X_n$ be n discrete random variables. The collection is specified by their joint pmf $p(x_1, x_2, \ldots, x_n)$ which satisfies:

1. $p(x_1, x_2, \ldots, x_n) \geq 0$
2. $\sum_{x_1, \ldots, x_n} p(x_1, \ldots, x_n) = 1$
3. $p(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = p(x_1, x_2, \ldots, x_n)$

Concepts such as marginal, conditional, etc. are defined as in the bivariate case.

Example 1

For 5 random variables

$$p(x_4, x_5) = \sum_{x_1\, x_2\, x_3} p(x_1, x_2, x_3, x_4, x_5)$$

$$p_{x_1, x_2, x_3 | x_4, x_5}(x_1, x_2, x_3) = \frac{p(x_1, x_2, x_3, x_4, x_5)}{p(x_4, x_5)}$$

Definition 1

The random variables $X_1, \ldots, X_n$ are **independent** if

$$p(x_1, x_2, \ldots, x_n) = p(x_1) \cdot p(x_2) \cdots p(x_n)$$

Multinomial Distribution

Context: Random experiment consist of n independent trials.

- The results of each trial can be classified in k classes.
- The probability that a trial generates a result in class 1, class 2 $, \ldots,$ class k is constant over the trials and these probabilities will be denoted by $p_1, p_2, \ldots, p_k$

The random variables $X_1, \cdots, X_k$ denote the number of trials that result in class 1, class 2, ..., class k, respectively, and they have multinomial joint pmf:

$$p\left(x_1, x_2, \ldots, x_k\right) = p\left(X_1 = x_1, \ldots, X_k = x_k\right) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

Note that: $x_1 + \cdots + x_4 = n \quad ; \quad p_1 + \ldots + p_k = 1.$

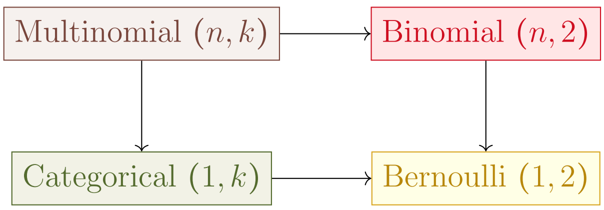
Example 2

If the allele of each of 10 independently obtained pea sections is determined and $p_1 = p(AA) = 0.2; p_2 = p\left(Aa\right) = 0.5; p_3 = p(aa) = 0.3$ and X_1 is the number of AA's, X_2 is the number of Aa's, X_3 is the number of aa's then

$$p\left(x_1, x_2, x_3\right) = \frac{10!}{x_1! x_2! x_3!} (0.2)^{x_1} (0.5)^{x_2} (0.3)^{x_3}$$

In particular,

$$p(2,5,3) = \frac{10!}{2!5!3!} (0.2)^2 (0.5)^5 (0.3)^3 = 0.085$$



Example 3

For the multinomial x_1, x_2, x_3, x_4 with $n = 15$ and $p_1 = 0,3, p_2 = 0.1$, $p_3 = 0.2$, and $p_4 = 0.4$ determine

a The marginals

(a) $p\left(x_1, x_2, x_3\right)$

(b) $p\left(x_1, x_2\right)$

(c) $p\left(x_1\right)$

b The conditionals

i $p\left(x_1, x_2, x_3 \mid x_4\right)$

ii $p\left(x_1, x_2 \mid x_3, x_4\right)$

iii $P\left(x_1 \mid x_2, x_3, x_4\right)$

Solution

We have

$$p(x_1, x_2, x_3, x_4) = \frac{15!}{x_1!x_2!x_3!x_4!}(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}(0.4)^{x_4}$$

a. i From the condition $x_1 + x_2 + x_3 + x_4 = 15$ we have $x_4 = 15 - x_1 - x_2 - x_3$ fixed

$$p(x_1, x_2, x_3) = \frac{15!}{x_1!x_2!x_3!(15 - x_4)!}(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}(0.4)^{(15-x_1-x_2-x_3)}$$

ii For $p(x_1, x_2)$ we have

$$p(x_1, x_2) = \sum_{x_3,x_4} \frac{15!}{x_1!x_2!x_3!x_4!}(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}(0.4)^{x_4}$$
$$= \sum_{x_3} \frac{15!}{x_1!x_2!x_3!(15 - x_1 - x_2 - x_3)!}(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}(0.4)^{15-x_1-x_2-x_3}$$

iii Finally for the single variable marginal

$$p(x_1) = \sum_{x_2,x_3,x_4} \frac{15!}{x_1!x_2 1x_3!x_4!}(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}(0.4)^{x_4}$$
$$= \frac{15!}{x_1(15 - x_1)!}(0.3)^{x_1}(0.7)^{15-x_1} \quad \leftarrow \quad \text{binomial}$$

b. i $p(x_1, x_2, x_3 \mid x_4)$

$$= \frac{p(x_1, x_2, x_3, x_4)}{p(x_4)}$$
$$= \frac{15!/(x_1!x_2!x_3!x_4!)}{15!/(x_1((15 - x_4)!)} \cdot \frac{(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}(0.4)^{x_4}}{(0.6)^{15-x_4}(0.4)^{x_4}}$$
$$= \frac{(15 - x_4)!}{x_1!x_2!x_3!} \cdot \frac{(0.3)^{x_1}(0.1)^{x_2}(0.2)^{x_3}}{(0.6)^{15-x_4}} \quad ; \quad x_1 + x_2 + x_3 = 15 - x_4$$

ii $p(x_1, x_2 \mid x_3, x_4) = \dfrac{p(x_1, x_2, x_3, x_4)}{p(x_3, x_4)}$

iii $p(x_1 \mid x_2, x_3, x_4)$

$$= \frac{p(x_1, x_2, x_3, x_4)}{p(x_2, x_3, x_4)} = 1 \text{ when } x_1 = 15 - x_2 - x_3 - x_4$$
$$= 0 \text{ otherwise.}$$