Probability and Stochastic Processes I - Lecture 8

Michael Evans
University of Toronto
http://www.utstat.utoronto.ca/mikevans/stac62/STAC622023.html

2023

II.5 Transformations and Change of Variables

- suppose we have a random vector $\mathbf{X} \in R^k$ and we transform this to be a new random vector $\mathbf{Y} = T(\mathbf{X}) \in R^l$

- given the probability distribution of $\mathbf{X}$, whether specified by a probability function $p_{\mathbf{X}}$ or a density $f_{\mathbf{X}}$, determine the probability function or density of $\mathbf{Y}$

discrete case

- so suppose $\mathbf{X}$ has prob. function $p_{\mathbf{X}}$ and now we want

$$p_{\mathbf{Y}}(\mathbf{y}) = P_{\mathbf{Y}}(\{\mathbf{y}\}) = P_{\mathbf{X}}(T^{-1}\{\mathbf{y}\}) = \sum_{\mathbf{x} \in T^{-1}\{\mathbf{y}\}} p_{\mathbf{X}}(\mathbf{x})$$

Example II.5.1 - suppose $p_{\mathbf{X}}(0,1) = 1/2, p_{\mathbf{X}}(1,0) = 1/3, p_{\mathbf{X}}(1,1) = 1/6$ and $y = T(x_1, x_2) = x_1 + x_2$

- so Y takes the values 1 and 2 (only) with positive probability and since $T^{-1}\{1\} = \{(0,1),(1,0)\}$ then $p_Y(1) = 1/2 + 1/3 = 5/6$ and since $T^{-1}\{2\} = \{(1,1)\}$ then $p_Y(2) = 1/6$

Example II.5.2 *Projections*

- suppose $k \geq 2$ and $(y_1, y_2) = T(x_1, \ldots, x_k) = (x_1, x_2)$ projection on the first two coordinates, then

$$T^{-1}\{\mathbf{y}\} = T^{-1}\{(y_1, y_2)'\} = \{(x_1, \ldots, x_k) : x_1 = y_1, x_2 = y_2\}$$

$$\begin{aligned} p_{\mathbf{Y}}(y_1, y_2) &= P_{\mathbf{X}}(T^{-1}\{\mathbf{y}\}) = \sum_{\mathbf{x} \in T^{-1}\{\mathbf{y}\}} p_{\mathbf{X}}(\mathbf{x}) \\ &= \sum_{(x_1, \ldots, x_k) : x_1 = y_1, x_2 = y_2} p_{\mathbf{X}}(x_1, \ldots, x_k) \\ &= \sum_{(x_3, \ldots, x_k) \in R^{k-2}} p_{\mathbf{X}}(y_1, y_2, x_3, \ldots, x_k) \end{aligned}$$

- also if $y = T(x_1, \ldots, x_k) = x_2$ projection on the second coordinate, then

$$T^{-1}\{y\} = T^{-1}\{y\} = \{(x_1, \ldots, x_k) : x_2 = y\}$$

$$\begin{aligned} p_Y(y) &= \sum_{(x_1, \ldots, x_k) : x_2 = y} p_{\mathbf{X}}(x_1, \ldots, x_k) \\ &= \sum_{(x_1, x_3, \ldots, x_k) \in R^{k-2}} p_{\mathbf{X}}(x_1, y, x_3, \ldots, x_k) \end{aligned}$$

- so the general approach for finding the probability functions of projections is to take the joint probability function and sum out all the remaining variables

Example II.5.3 *Multinomial*$(n, p_1, \ldots, p_k)$ *distribution*

- when let $\mathbf{X} = (X_1, \ldots, X_k)' \sim multinomial(n, p_1, \ldots, p_k)$ then $p_{\mathbf{X}}$ is only positive on $\mathbf{a} \in R^k$ when

$$a_i \in \{0, \ldots, n\} \text{ and } a_1 + \cdots + a_k = n \qquad (*)$$

and has

$$p_{\mathbf{X}}(\mathbf{a}) = \binom{n}{a_1\ a_2\ \ldots\ a_k} p_1^{a_1} \cdots p_k^{a_k}$$

- suppose $k \geq 2$ and $(y_1, y_2) = T(x_1, \ldots, x_k) = (x_1, x_2)$ so we want the distribution of $\mathbf{Y} = (X_1, X_2)'$

- by (*) $y_1, y_2, a_3, \ldots, a_k \in \{0, \ldots, n\}$ and $y_1 + y_2 + a_3 + \cdots + a_k = n$ iff

$$a_3, \ldots, a_k \in \{0, \ldots, n - y_1 - y_2\} \text{ and } a_3 + \cdots + a_k = n - y_1 - y_2 \qquad (**)$$

- therefore

$$p_Y(y_1, y_2) = \sum_{(a_3,\ldots,a_k) \text{ sat. } (**)} \binom{n}{y_1\ y_2\ a_3 \ldots\ a_k} p_1^{y_1} p_2^{y_2} p_3^{a_3} \cdots p_k^{a_k}$$

$$= \frac{n!}{y!\, y_2!(n - y_1 - y_2)!} p_1^{y_1} p_2^{y_2} \sum_{(a_3,\ldots,a_k) \text{ sat. } (**)} \frac{(n - y_1 - y_2)!}{a_3! \cdots a_k!} p_3^{a_3} \cdots p_k^{a_k}$$

$$= \binom{n}{y_1\ y_2\ n - y_1 - y_2} p_1^{y_1} p_2^{y_2} (1 - p_1 - p_2)^{n - y_1 - y_2} \times$$

$$\sum_{(a_3,\ldots,a_k) \text{ sat.g } (**)} \binom{n - y_1 - y_2}{a_3 \ldots a_k} \left(\frac{p_3}{1 - p_1 - p_2}\right)^{a_3} \cdots \left(\frac{p_k}{1 - p_1 - p_2}\right)^{a_k}$$

$$= \binom{n}{y_1\ y_2\ n - y_1 - y_2} p_1^{y_1} p_2^{y_2} (1 - p_1 - p_2)^{n - y_1 - y_2}$$

since the second term is the sum of all

$$\text{multinomial}\left(n - y_1 - y_2, \frac{p_3}{1 - p_1 - p_2}, \ldots, \frac{p_k}{1 - p_1 - p_2}\right)$$

probabilities and so $(X_1, X_2) \sim \text{multinomial}(n, p_1, p_2, 1 - p_1 - p_2)$ ■

Exercise II.5.1 If $\mathbf{X} = (X_1, \ldots, X_k)' \sim \text{multinomial}(n, p_1, \ldots, p_k)$ then prove $X_i \sim \text{binomial}(n, p_i) = \text{multinomial}(n, p_i, 1 - p_i)$.

note - these results for the multinomial are easy to see intuitively since the multinomial arises by categorizing n independent observations into k mutually disjoint categories and when we project onto l coordinates we are now categorizing into $l + 1$ mutually disjoint categories

Exercise II.5.2 Use the above note to determine the distribution of $Y = X_1 + \cdots + X_l$ for $l \leq k$ when

$$(X_1, \ldots, X_k)' \sim \text{multinomial}(n, p_1, \ldots, p_k).$$

note - in the discrete case if T is 1-1 ($T(\mathbf{x}_1) = T(\mathbf{x}_2)$ iff $\mathbf{x}_1 = \mathbf{x}_2$) then

$$p_{\mathbf{Y}}(\mathbf{y}) = P_{\mathbf{X}}(T^{-1}\{\mathbf{y}\} = p_{\mathbf{X}}(T^{-1}\{\mathbf{y}\})$$

whenever $T^{-1}\{\mathbf{y}\} \neq \phi$

- depending on the transformation T it could be that $\mathbf{Y} = T(\mathbf{X})$ has a discrete distribution no matter what kind of distribution $\mathbf{X}$ has

Example II.5.4

- suppose $T(\mathbf{x}) = \mathbf{c} \in R^l$ for every $\mathbf{x}$, then

$$p_{\mathbf{Y}}(\mathbf{y}) = P_{\mathbf{X}}(T^{-1}\{\mathbf{y}\}) = \begin{cases} P_{\mathbf{X}}(R^k) = 1 & \text{if } \mathbf{y} = \mathbf{c} \\ P_{\mathbf{X}}(\phi) = 0 & \text{if } \mathbf{y} \neq \mathbf{c} \end{cases}$$

and the distribution of $\mathbf{Y}$ is *degenerate* at $\mathbf{c}$

- suppose $X \sim N(0,1)$ so $P(X \leq 0) = P(X > 0) = 1/2$, then if

$$Y = T(X) = I_{(-\infty,0]}(X) = \begin{cases} 1 & \text{if } X \leq 0 \\ 0 & \text{if } X > 0 \end{cases}$$

then

$$p(1) = P(X \leq 0) = 1/2 \text{ and } p(1) = P(X > 0) = 1/2$$

so $Y \sim \text{Bernoulli}(1/2)$ ■

Definition II.5.1 For $A \subset \Omega$ the function $I_A : \Omega \to R^1$ given by

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \in A^c \end{cases}$$

is called the *indicator function* of A. ■

Exercise II.5.3 If $(\Omega, \mathcal{A}, P)$ is a probability model and $A \in \mathcal{A}$ then $Y = I_A$ is a random variable with $Y \sim \text{Bernoulli}(P(A))$.