STAT 330: Lecture 10

2024 Spring

June 13, 2024

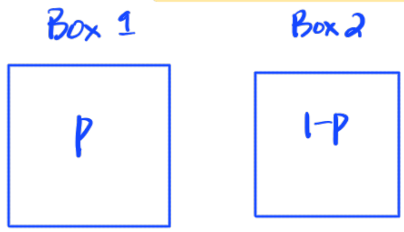
In the past few lectures We have learned the basics of joint distribution.

This Lecture We will learn two important multivariate random variables.

$\begin{cases} \text{Multinomial} \\ \text{Bivariate Normal} \end{cases}$

Multinomial Distribution

Examples: If we have two boxes, both having many balls. The probability of choosing a ball from Box 1 is p. Suppose we choose n times and get n balls, and denote the number of balls from Box 1 as X. What is the distribution of X.



$X \sim Bin(n, p)$

If we have $k+1$ boxes, all have many balls. The probability of choosing a ball from Box i is p_i. After n draws, let's denote the number of balls from Box i as X_i.

① $\sum_{i=1}^{K+1} P_i = 1$, $P_{K+1} = 1 - \sum_{i=1}^{K} P_i$

② $X_{K+1} = n - \sum_{i=1}^{K} X_i$

need parameter: $P_1 \cdots P_K$

random variable: $X_1 \cdots X_K$

How to characterize the dist'n of $(X_1, \cdots, X_K)$?

Definition:

$(X_1, \cdots, X_K)$ are joint discrete rvs. with pmf

$$f(x_1, \cdots, x_K) = Pr(X_1 = x_1, \cdots, X_K = x_K) = \frac{n!}{x_1! \cdots x_K!\, x_{K+1}!} \cdot P_1^{x_1} \cdots P_K^{x_K} \cdot P_{K+1}^{x_{K+1}}$$

$P_{K+1} = 1 - \sum_{i=1}^{K} P_i$, $x_{K+1} = n - \sum_{i=1}^{K} x_i$

Support: $x_i = 0, 1, \cdots, n$ for $i = 1, \cdots, K+1$.

$P_i \in (0,1)$, $i = 1, \cdots, K+1$

$\boxed{\text{Multinomial Distribution}}$

$Mult(n, P_1, \cdots P_K)$

If $K=1$, reduced to binomial.

PMF Explanation:

$\frac{n!}{x_1! \cdots x_K!\, x_{K+1}!}$ = # of combinations s.t. we have x_i balls from box i.

$P_1^{x_1} \cdots P_K^{x_K} P_{K+1}^{x_{K+1}}$ = probability of one combination.

Properties: $(X_1 \cdots X_K) \sim Mult(n, p_1, \cdots p_K)$

1. Joint mgf.

$$M(t_1, \cdots, t_K) = E\{e^{t_1X_1 + \cdots + t_KX_K}\} = \underline{(p_1e^{t_1} + \cdots + p_Ke^{t_K} + p_{K+1})^n}$$

$p_{K+1} = 1 - p_1 - \cdots - p_K$

verify by yourself, by creating a multinomial pmf (by scaling).

2. Marginal Dist'n.

① From joint mgf → marginal mgf.

$$M_{X_i}(t_i) = M(0, \cdots 0, t_i, 0 \cdots, 0) = (p_1 + \cdots + p_{i-1} + p_ie^{t_i} + p_{i+1} + \cdots + p_{K+1})^n$$
$$= (p_ie^{t_i} + 1 - p_i)^n \text{ Binomial mgf.}$$

$\Rightarrow X_i \sim Bin(n, p_i)$

② From joint mgf → joint mgf. (a subset of $(X_1 \cdots X_n)$)

e.g. $M_{X_1,X_2}(t_1, t_2) = M(t_1, t_2, 0, \cdots, 0) = (p_1e^{t_1} + p_2e^{t_2} + \underbrace{p_3 + \cdots + p_{K+1}}_{=1-p_1-p_2})^n$

$(X_1, X_2) \sim Mult(n, p_1, p_2)$

3. $T = X_i + X_j$, $i \neq j$. what is the dist'n of T

Intuitively, success: a ball is from box i or box j.
failure: otherwise.

T = # of successes out of n draws.

$T \sim Bin(n, p_i + p_j)$

$$M_T(t) = E(e^{tT}) = E(e^{tX_1 + tX_2}) = M(t, t, 0, \cdots, 0)$$
$$= [(p_1 + p_2)e^t + 1 - p_1 - p_2]^n \text{ Binomial mgf}$$

$T \sim Bin(n, p_1 + p_2)$

4. Moments

$E(X_i) = np_i$, $Var(X_i) = np_i(1-p_i)$

$Cov(X_i, X_j) = ?$

$Var(X_i + X_j) = Var(X_i) + Var(X_j) + 2Cov(X_i, X_j)$

$n(p_i+p_j)(1-p_i-p_j)$

$\Rightarrow Cov(X_i, X_j) = -np_ip_j$ by solving the equation above.

5. Conditional distributions

Find $X_i \mid X_j = x_j$

$$f(x_i|x_j) = \frac{f(x_i, x_j)}{f(x_j)} = \frac{\frac{\cancel{n!}}{x_i!\,\cancel{x_j!}\,(n-x_i-x_j)!} \cdot p_i^{x_i} \cancel{p_j^{x_j}} (1-p_i-p_j)^{n-x_i-x_j}}{\frac{\cancel{n!}}{\cancel{x_j!}\,(n-x_j)!} \cdot \cancel{p_j^{x_j}} (1-p_j)^{n-x_j}}$$

$$= \frac{(n-x_j)!}{x_i!\,(n-x_i-x_j)!} \cdot \left(\frac{p_i}{1-p_j}\right)^{x_i} \left(\frac{1-p_i-p_j}{1-p_j}\right)^{n-x_i-x_j}$$

$$\sim Bin\left(n-x_j, \frac{p_i}{1-p_j}\right)$$

Verify $X_i \mid X_i + X_j = t \sim Bin\left(t, \frac{p_i}{p_i+p_j}\right)$ by yourself.

$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

Bivariate Normal Distribution We say $\boldsymbol{X}^{\mathrm{T}} = (X_1, X_2)$ follows a bivariate normal distribution if its joint pdf is

$$f(x_1, x_2) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}, \quad |x_1|, |x_2| < \infty.$$

$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$, Σ is positive definite.

$\boxed{\text{For any } x \neq 0,\ x^T \Sigma x > 0 \quad \in \mathbb{R}.}$

$|\Sigma|$: determinant of $\Sigma > 0$

$|\Sigma| = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = (1-\rho^2)\sigma_1^2\sigma_2^2$

Or, the joint pdf can be given by

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - 2\frac{\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right\}\right]$$

Properties:

$t = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$, $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

1. Joint mgf.

   $M(t_1, t_2) = E[e^{t_1X_1 + t_2X_2}] = \exp\left(t^T\mu + \frac{1}{2}t^T\Sigma t\right)$

   $= \exp\left(t_1\mu_1 + t_2\mu_2 + \frac{1}{2}t_1^2\sigma_1^2 + \frac{1}{2}t_2^2\sigma_2^2 + t_1t_2\rho\sigma_1\sigma_2\right)$

   $t_1, t_2 \in \mathbb{R}$.

2. Marginal

   $M_{X_1}(t_1) = M(t_1, 0) = \exp\left(t_1\mu_1 + \frac{1}{2}t_1^2\sigma_1^2\right)$ mgf of normal distn.

   $X_1 \sim \mathrm{Norm}(\mu_1, \sigma_1^2)$

   Similarly. $X_2 \sim \mathrm{Norm}(\mu_2, \sigma_2^2)$ (using $M_{X_2}(t_2) = M(0, t_2)$)

   $\begin{cases} E(X_1) = \mu_1 \\ E(X_2) = \mu_2 \end{cases}$

3. conditional dist'n.

$X_2 | X_1 = x_1 \sim \text{Norm}(\mu_2 + \rho\sigma_2(x_1 - \mu_1)/\sigma_1,\ \sigma_2^2(1-\rho^2))$

$X_1 | X_2 = x_2 \sim \text{Norm}(\mu_1 + \rho\sigma_1(x_2 - \mu_2)/\sigma_2,\ \sigma_1^2(1-\rho^2))$

verify one of them by yourself. $\boxed{f(x_2|x_1) = \frac{f(x_1, x_2)}{f(x_1)}}$

4. Find $\text{Cov}(X_1, X_2)$

$\text{Cov}(X_1, X_2) = E(X_1 X_2) - E(X_1) \cdot E(X_2)$, $E(X_1) = \mu_1$, $E(X_2) = \mu_2$

$E(X_1 X_2) = E\{E(X_1 X_2 | X_1)\}$

$E(X_1 X_2 | X_1 = x_1) = x_1 E(X_2 | X_1 = x_1) = h(x_1)$

change x_1 to X_1.

$= E\{X_1 E(X_2 | X_1)\}$

$h(X_1) = E(X_1 X_2 | X_1) = X_1 E(X_2 | X_1)$

$= E[X_1\{\mu_2 + \rho\sigma_2(X_1 - \mu_1)/\sigma_1\}]$

$= \mu_2 E(X_1) + \rho\frac{\sigma_2}{\sigma_1} E(X_1^2) - \frac{\rho\sigma_2}{\sigma_1}\mu_1 \cdot E(X_1)$

$= \rho\sigma_1\sigma_2 + \mu_1\mu_2$

$\Rightarrow \text{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2$

Meanings of the Parameters:

$X \sim \text{BVN}(\mu, \Sigma)$.

$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} E(X_1) \\ E(X_2) \end{bmatrix}$ mean vector

$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_1, X_2) & \text{Var}(X_2) \end{bmatrix}$

$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\ \text{Var}(X_2)}} = \frac{\rho\sigma_1\sigma_2}{\sqrt{\sigma_1^2}\sqrt{\sigma_2^2}} = \rho$

variance-covariance matrix. / covariance matrix / variance matrix