H. Two Special Multivariate Distributions

1. Multinomial distribution
   a. Extension of binomial
   b. Result of categorizing \( n \) independent items into one of \( K \) categories.
   c. \( X_i \) is count in cell \( i \), for \( i \leq K \)
   d. \( \pi_i \) = probability for category \( i \), \( i \leq K \).
   e. \( P(X_j = x_j \forall j) = \left(\frac{n}{x_1 x_2 \cdots x_K}\right) \prod_{i=1}^{K} \pi_i^{x_i} \), because:
      i. If items are uniquely identifiable as \( 1, \ldots, n \),
      ii. Let \( U_1, \ldots, U_n \) be the indicators for which category the items are in.
   f. \( U_j \in \{1, \ldots, K\} \) for each \( j \)
   g. Let \( x_i \) be the number of \( u_j \) that equal \( i \), for \( i \in \{1, \ldots, K\} \).
   h. \( P(U_j = u_j \forall j) = \prod_{i=1}^{K} \pi_{u_i} = \prod_{i=1}^{K} \pi_i^{x_i} \), because:
   i. While this is the probability of a specific sequence of categories, it depends only on the number in each group, and not on the order they came in.
   j. The number of orderings consistent with these counts is \( \frac{n!}{x_1! \cdots x_K!} \).
   k. \( P(X) = \frac{n!}{x_1! \cdots x_K!} \pi^{x} \)
      • for \( x_i \geq 0 \forall i \), \( x_K = n - \sum_{i=1}^{K-1} x_i \),
      • \( \pi_i \in [0,1] \forall i \), \( \pi_K = 1 - \pi_1 - \cdots - \pi_{K-1} \).
   l. Stays multinomial after rearranging categories
   m. Stays multinomial after collapsing categories

2. Bivariate Normal Distribution: Special case with standard normal marginals.
   a. \( Y_1 \sim N(0,1) \), \( Y_2 | Y_1 = N(\rho Y_1, 1 - \rho^2) \)
      i. Implies \( E(Y_2 | Y_1) = \rho Y_1 \)
   b. Joint probability density function
      \[ f_{Y_1,Y_2}(y_1, y_2) = \frac{\exp(-\frac{1}{2}(y_1^2 - 2\rho y_1 y_2 + y_2^2)/(2(1-\rho^2)))}{2\pi \sqrt{1-\rho^2}}. \]
      i. The probability density function is product of marginal and conditional probability density functions
      \[ f_{Y_1,Y_2}(y_1, y_2) = \frac{\exp(-\frac{y_1^2}{2})/\sqrt{2\pi}}{\exp(-\frac{y_2^2}{2})/\sqrt{2\pi}} \]
      ii. Combining the arguments to the exponential function and simplifying gives the above.
c. Argument is symmetric if you change direction of conditioning:
\[ f_{Y_1,Y_2}(y_1, y_2) = \frac{\exp(-y_1^2/2) \exp(-y_2^2/(2(1-\rho^2)))}{\sqrt{2\pi(1-\rho^2)}} \]

   i. Hence \( Y_2 \sim N(0,1) \), \( Y_1|Y_2 \sim N(\rho Y_2, 1-\rho^2) \)

   d. Also \( \text{Cov}(Y_1, Y_2) = \rho \).
   
   i. Because \( \text{Cov}(Y_1, Y_2) = E(Y_1 Y_2) = E(E(Y_2|Y_1)) = E(E(Y_1|Y_2)) = E(\rho Y_2^2) = \rho \).
   
   e. \( \rho \) is also the correlation.

   ii. Because the marginal standard deviations are 1.

4. More on the Multivariate Normal
a. Sums of components of bivariate normals are still normal.
   
   i. If \( (Y_1, Y_2) \) bivariate normal, \( a_1 \) and \( a_2 \) constants not both 0, then \( a_1 Y_1 + a_2 Y_2 \) normal.

   ii. Demonstrate using change of variables formula for probability density function

   • WOLOG \( a_2 \neq 0 \)
   
   • \( Z_1 = a_1 Y_1 + a_2 Y_2 \), \( Z_2 = Y_1 \)
   
   • \( Y_1 = Z_1 / a_2, Y_2 = (Z_1 - a_2 Z_2)/a_2 \).

   iii. Jacobian is a constant

   iv. Quantity in exponential is

   \[ -\frac{(z_2 - \mu_1)^2}{2\sigma_1^2} + \rho \frac{(z_2 - \mu_1)(z_1 - a_2 \sigma_2)}{a_2 \sigma_2} - \frac{(z_1 - a_2 \sigma_2)^2}{2a_2^2 \sigma_2^2} \]

   • Multipliers of \( z_1^2 \) and \( z_2^2 \) are negative.

   • Can check that cross term keeps \( \rho \) in \([-1,1]\).

   v. Marginalize to get distribution of \( Z_1 \).

b. Calculate in R using `dmvnorm(x,c(0,0),diag(c(1,1)))`.
   
   i. Must have previously loaded `library(mvtnorm)`
   
   ii. There are also options for quantile, distribution function.

3. General Form of Bivariate Normal
   a. Here \( Y_1 \sim N(\mu_1, \sigma_1) \), \( Y_2|Y_1 \sim N(\mu_2 + \sigma_2 \rho (Y_1 - \mu_1)/\sigma_1, \sigma_2^2(1-\rho^2)) \).

   b. Probability density function

   \[ f_{Y_1,Y_2}(y_1, y_2) = \frac{\exp\left(-\frac{(y_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(y_1 - \mu_1)(y_2 - \mu_2)}{\sigma_1 \sigma_2} + \frac{(y_2 - \mu_2)^2}{\sigma_2^2}\right)}{2\sigma_1 \sigma_2 \pi \sqrt{1-\rho^2}} \]

   i. Five parameters:

   • Two expectations

   • Two marginal standard deviations

   • One correlation.

   ii. Extreme example: \( X \sim N(0,1) \), \( Y = X \) bivariate normal with \( \rho = 1 \):

   • Doesn’t have a probability density function in two dimensions.

   • Bivariate normal, but a degenerate case.

5. Correlation 0 implies independence for multivariate normals
   a. Always, independence implies zero correlation

   b. Bivariate normality is needed for the opposite direction.

   i. Probability density function

   \[ f_{Y_1,Y_2}(y_1, y_2) = \frac{\exp\left(-\frac{(y_1 - \mu_1)^2}{\sigma_1^2} + \frac{(y_2 - \mu_2)^2}{\sigma_2^2}\right)}{2\pi \sigma_1 \sigma_2 \sqrt{1-\rho^2}} \]

   ii. Setting \( \rho = 0 \) gives

   \[ f_{Y_1,Y_2}(y_1, y_2) = \exp\left(-\frac{(y_1 - \mu_1)^2}{\sigma_1^2} + \frac{(y_2 - \mu_2)^2}{\sigma_2^2}\right) \]

   c. Multivariate normal not just a joint distribution with normal marginals.

   i. \( X \sim N(0,1) \), \( Y = \begin{cases} X & \text{if } |X| \leq c \\ -X & \text{otherwise} \end{cases} \)

   ii. Doesn’t have a probability density function in two dimensions, and not even a degenerate case of bivariate normal.

   iii. Can pick \( c \) to give zero correlation.

   iv. Not independent for any \( c \).

   v. \( X - Y \) clearly not normal. See Fig. 41.
Fig. 40: Univariate Normal Data that are Not Bivariate Normal