Annette J. Dobson

An Introduction to Generalized Linear Models

Solutions

by Konstantinos Kirillov

CHAPTER 1

Exercise 1.1:

Let Y_1 and Y_2 be independent random variables with $Y_1 \sim N(1, 3)$ and $Y_2 \sim N(2, 5)$. If $W_1 = Y_1 + 2Y_2$ and $W_2 = 4Y_1 - Y_2$ what is the joint distribution of W_1 and W_2?

SOLUTION:

A reminder from the book:

1.4.1 Normal distributions:

1. If the random variable Y has the Normal distribution with mean μ and variance σ^2, its probability density function is:

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma^2}\right)^2\right]$$

We denote this by $Y \sim N(\mu, \sigma^2)$.

2. The Normal distribution with $\mu = 0$ and $\sigma^2 = 1$, $Y \sim N(0, 1)$, is called the **standard Normal distribution**.

3. Let $Y_1, \dots, Y_n$ denote Normally distributed random variables with $Y_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, \dots, n$ and let the covariance of Y_i and Y_j be denoted by:

$$cov(Y_i, Y_j) = \rho_{ij}\sigma_i\sigma_j,$$

where ρ_{ij} is the correlation coefficient for Y_i and Y_j. Then the joint distribution of the Y_i's is the **multivariate Normal distribution** with mean vector $\mu = [\mu_1, \dots, \mu_n]^T$ and variance-covariance matrix V with diagonal elements σ_i^2 and non-diagonal elements $\rho_{ij}\sigma_i\sigma_j$ for $i \neq j$. We write this as:

$$\boldsymbol{y} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{V}), \text{ where } \boldsymbol{y} = [Y_1, \dots, Y_n]^T$$

4. Suppose the random variables $Y_1, \dots, Y_n$ are independent and normally distributed with the distributions $Y_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, \dots, n$. If

$$W = a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n,$$

where the a_i's are constants. Then W is also Normally distributed, so that:

$$W = \sum_{i=1}^{n} a_i Y_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

It seems that the joint distribution of two normally distributed variables is yet another normal distribution. In this exercise, in order to find the joint distribution of W_1 and W_2, we first need to determine the mean, the variance and the covariance of W_1 and W_2 and then use those to derive the joint distribution.

<u>Given that:</u>
$Y_1 \sim N(1, 3)$
$Y_2 \sim N(2, 5)$

First, let us find the means of W_1 and W_2:

$E(W_1) = E(Y_1 + 2Y_2) = E(Y_1) + 2E(Y_2) = 1 + 2 \cdot 2 = 5 \Rightarrow \boldsymbol{E(W_1) = 5}$

$E(W_2) = E(4Y_1 - Y_2) = 4E(Y_1) - E(Y_2) = 4 \cdot 1 - 2 = 2 \Rightarrow \boldsymbol{E(W_2) = 2}$

Next, let us calculate the variances of W_1 and W_2:

$Var(W_1) = Var(Y_1 + 2Y_2) = Var(Y_1) + 2^2 \cdot Var(Y_2) = 3 + 4 \cdot 5 = 23 \Rightarrow \boldsymbol{Var(W_1) = 23}$

$Var(W_2) = Var(4Y_1 - Y_2) = 4^2 \cdot Var(Y_1) + Var(Y_2) = 16 \cdot 3 + 5 = 53 \Rightarrow \boldsymbol{Var(W_2) = 53}$

And finally, let us also compute the covariance between W_1 and W_2:

$Cov(W_1, W_2) = Cov(Y_1 + 2Y_2, 4Y_1 - Y_2) =$
$= Cov(Y_1, 4Y_1) + Cov(Y_1, -Y_2) + Cov(2Y_2, 4Y_1) + Cov(2Y_2, -Y_2) =$
$= 4Var(Y_1) - Cov(Y_1, Y_2) + 8Cov(Y_2, Y_1) - 2Var(Y_2) =$
$= 4 \cdot 3 - 0 + 8 \cdot 0 - 2 \cdot 5 = 2 \Rightarrow$
$\Rightarrow \boldsymbol{Cov(W_1, W_2) = 2}$

Therefore, the joint distribution will be:

$$\begin{pmatrix} \boldsymbol{W_1} \\ \boldsymbol{W_2} \end{pmatrix} \sim \boldsymbol{N}\left[\begin{pmatrix} \mathbf{5} \\ \mathbf{2} \end{pmatrix}, \begin{pmatrix} \mathbf{23} & \mathbf{2} \\ \mathbf{2} & \mathbf{53} \end{pmatrix}\right]$$

The correlation coefficient between W_1 and W_2 in this case shall be:

$$\rho = \frac{Cov(W_1, W_2)}{\sigma_{W_1} \cdot \sigma_{W_2}} = \frac{Cov(W_1, W_2)}{\sqrt{Var(W_1)} \cdot \sqrt{Var(W_2)}} = \frac{2}{\sqrt{23} \cdot \sqrt{53}} \approx \frac{2}{4.8 \cdot 7.3} \approx 0.057 \Rightarrow \boldsymbol{\rho = 0.057}$$

Therefore, another way to express the joint distribution, would be:

$$f(W_1, W_2) = \frac{1}{2 \cdot \pi \cdot \sigma_{W_1} \cdot \sigma_{W_2} \cdot \sqrt{1 - \rho^2}} exp\left[-\frac{Z_{W_1}^2 - 2 \cdot Z_{W_1} \cdot Z_{W_2} + Z_{W_2}^2}{2 \cdot \sqrt{1 - \rho^2}}\right] \Rightarrow$$

$$\Rightarrow \boldsymbol{f(W_1, W_2)} = \frac{\mathbf{1}}{\mathbf{2} \cdot \boldsymbol{\pi} \cdot \mathbf{4.8} \cdot \mathbf{7.3} \cdot \sqrt{\mathbf{1 - 0.057^2}}} \boldsymbol{exp}\left[-\frac{\boldsymbol{Z_{W_1}^2 - 2 \cdot Z_{W_1} \cdot Z_{W_2} + Z_{W_2}^2}}{\mathbf{2} \cdot \sqrt{\mathbf{1 - 0.057^2}}}\right]$$

Where:

$$Z_{W_1} = \frac{W_1 - \mu_{W_1}}{\sigma_{W_1}}$$

$$Z_{W_2} = \frac{W_2 - \mu_{W_2}}{\sigma_{W_2}}$$

Exercise 1.2:

Let $\boldsymbol{Y_1}$ and $\boldsymbol{Y_2}$ be independent random variables with $\boldsymbol{Y_1 \sim N(0,1)}$ and $\boldsymbol{Y_2 \sim N(3,4)}$.
a. What is the distribution of $\boldsymbol{Y_1^2}$?
b. If $\boldsymbol{y = \begin{bmatrix} Y_1 \\ (Y_2 - 3)/2 \end{bmatrix}}$, obtain an expression for $\boldsymbol{y^T y}$. What is its distribution?
c. If $\boldsymbol{y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}}$ and its distribution is $\boldsymbol{y \sim N(\mu, V)}$, obtain an expression for $\boldsymbol{y^T V^{-1} y}$. What is its distribution?

SOLUTION:

A reminder from the book:

1.4.2 Chi-squared distribution:

1. The **central chi-squared distribution** with n degrees of freedom is defined as the sum of squares of n independent random variables $Z_1, \dots, Z_n$ each with the standard Normal distribution. It is denoted by:

$$X^2 = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$$

In matrix notation, if $\mathbf{z} = [Z_1, \dots, Z_n]^T$ then $\mathbf{z}^T\mathbf{z} = \sum_{i=1}^{n} Z_i^2$ so that $X^2 = \mathbf{z}^T\mathbf{z} \sim \chi^2(n)$.

2. If X^2 has the distribution $\chi^2(n)$, then its expected value is $E(X^2) = n$ and its variance is $Var(X^2) = 2n$.

3. If $Y_1, \dots, Y_n$ are independent Normally distributed random variables each with the distribution $Y_i \sim N(\mu_i, \sigma_i^2)$ then:

$$X^2 = \sum_{i=1}^{n} \left(\frac{Y_i - \mu_i}{\sigma_i}\right)^2 \sim \chi^2(n)$$

because each of the variables $Z_i = (Y_i - \mu_i)/\sigma_i$ has the standard Normal distribution $N(0,1)$.

4. Let $Z_1, \dots, Z_n$ be independent random variables each with the distribution $N(0,1)$ and let $Y_i = Z_i + \mu_i$, where at least one of the μ_i's is non-zero. Then the distribution of:

$$\sum Y_i^2 = \sum (Z_i + \mu_i)^2 = \sum Z_i^2 + 2\sum Z_i\mu_i + \sum \mu_i^2$$

has larger mean $n + \lambda$ and larger variance $2n + 4\lambda$ than $\chi^2(n)$ where $\lambda = \sum \mu_i^2$. This is called the **non-central chi-squared distribution** with n degrees of freedom and **non-centrality parameter** λ. It is denoted by $\chi^2(n, \lambda)$.

5. Suppose that the Y_i's are not necessarily independent and the vector $\boldsymbol{y} = [Y_1, \dots, Y_n]^T$ has the multivariate normal distribution $\boldsymbol{y} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{V})$ where the variance-covariance matrix $\boldsymbol{V}$ is non-singular and its inverse is $\boldsymbol{V}^{-1}$. Then:

$$X^2 = (\boldsymbol{y} - \boldsymbol{\mu})^T \boldsymbol{V}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}) \sim \chi^2(n)$$

6. More generally if $\boldsymbol{y} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{V})$ then the random variable $\boldsymbol{y}^T \boldsymbol{V}^{-1} \boldsymbol{y}$ has the non-central chi-squared distribution $\chi^2(n, \lambda)$ where $\lambda = \boldsymbol{\mu}^T \boldsymbol{V}^{-1} \boldsymbol{\mu}$.

7. If $X_1^2, \dots, X_m^2$ are m independent random variables with the chi-squared distributions $X_i^2 \sim \chi^2(n_i, \lambda_i)$, which may or may not be central, then their sum also has a chi-squared distribution with $\sum n_i$ degrees of freedom and non-centrality parameter $\sum \lambda_i$, i.e.,

$$\sum_{i=1}^{m} X_i^2 \sim \chi^2\left(\sum_{i=1}^{m} n_i, \sum_{i=1}^{m} \lambda_i\right)$$

This is called the **reproductive property** of the chi-squared distribution.

8. Let $\boldsymbol{y} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{V})$, where $\boldsymbol{y}$ has n elements but the Y_i's are not independent so that $\boldsymbol{V}$ is singular with rank $k < n$ and the inverse of $\boldsymbol{V}$ is not uniquely defined. Let $\boldsymbol{V}^-$ denote a generalized inverse of $\boldsymbol{V}$. Then the random variable $\boldsymbol{y}^T \boldsymbol{V}^- \boldsymbol{y}$ has the non-central chi-squared distribution with k degrees of freedom and non-centrality parameter $\lambda = \boldsymbol{\mu}^T \boldsymbol{V}^- \boldsymbol{\mu}$.

a. As property 1 from above would suggest,he chi-squared distribution with n degrees of freedom, $\chi^2(n)$, is the distribution of the sum of the squares of n independent standard normal random variables. If Y_1 is a random variable following a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$ ($Y_1 \sim N(0,1)$), then the distribution of Y_1^2 is a special case of the **chi-squared distribution with one degree of freedom, $\boldsymbol{\chi^2(1)}$**. Meaning that:

$$\boldsymbol{Y_1^2 \sim \chi^2(1)}$$

The chi-squared distribution with 1 degree of freedom is sometimes referred to as the exponential distribution with rate parameter $\lambda = 2$ ($mean = 1/\lambda = 1/2$, $variance = 1/\lambda^2 = 1/4$).

So, the distribution of Y_1^2 is $\chi^2(1)$ or equivalently, an exponential distribution with rate parameter $\lambda = 2$.

b. The expression $y^T y$ is the dot product of the vector y with itself. So:

$$y^T y = \begin{bmatrix} Y_1 & \frac{Y_2 - 3}{2} \end{bmatrix} \begin{bmatrix} Y_1 \\ \frac{Y_2 - 3}{2} \end{bmatrix} = Y_1^2 + \left(\frac{Y_2 - 3}{2}\right)^2 \Rightarrow \boldsymbol{y^T y = Y_1^2 + \frac{Y_2^2 - 6 \cdot Y_2 + 9}{4}}$$

We know that $Y_1 \sim N(0,1)$ and $Y_2 \sim N(3,4)$, and that they are independent. We also know (form a) that Y_1^2 is a special case of the chi-squared distribution with one degree of freedom, $\chi^2(1)$, or in other words: $Y_1^2 \sim \chi^2(1)$.

Furthermore, we are given that: $Y_2 \sim N(3,4)$, thus:

$$Y_2 \sim N(3,4) \Rightarrow Y_2 - 3 \sim N(0,4) \Rightarrow \frac{Y_2 - 3}{2} \sim N(0,1) \Rightarrow \left(\frac{Y_2 - 3}{2}\right)^2 \sim \chi^2(1)$$

Since both Y_1^2 and $\left(\frac{Y_2-3}{2}\right)^2$ are independent and follow a chi-squared distribution with 1 degree of freedom, then it follows that their sum will also follow the chi-squared distribution, but with two degrees of freedom, that are coming from the two terms combined. Therefore:

$$y^T y = Y_1^2 + \frac{Y_2^2 - 6 \cdot Y_2 + 9}{4} \sim \chi^2(2)$$

c.