Annette J. Dobson

An Introduction to Generalized Linear Models

Solutions

by Konstantinos Kirillov

**CHAPTER 1**

***Exercise 1.1:***

**Let and be independent random variables with**  **and .   
If**  **and what is the joint distribution of and ?**

***SOLUTION:***

*A reminder from the book:*

1.4.1 Normal distributions:

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

We denote this by .

1. The Normal distribution with and , , is called the **standard Normal distribution**.
2. Let denote Normally distributed random variables with  for and let the covariance of and be denoted by:

where is the correlation coefficient for and . Then the joint distribution of the ’s is the **multivariate Normal distribution** with mean vector and variance-covariance matrix with diagonal elements and non-diagonal elements for . We write this as:

, where

1. Suppose the random variables are independent and normally distributed with the distributions for . If

where the ’s are constants. Then is also Normally distributed, so that:

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

Given that:

First, let us find the means of and :

Next, let us calculate the variances of and :

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

Therefore, the joint distribution will be:

The correlation coefficient between and in this case shall be:

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

Where:

***Exercise 1.2:***

**Let and be independent random variables with and .   
a. What is the distribution of ?  
b.** **If , obtain an expression for . What is its distribution?  
c. If and its distribution is , obtain an expression for . What is its distribution?**

***SOLUTION:***

*A reminder from the book:*

1.4.2 Chi-squared distribution:

1. The **central chi-squared distribution** with degrees of freedom is defined as the sum of squares of independent random variables each with the standard Normal distri-bution. It is denoted by:

In matrix notation, if then so that .

1. If has the distribution , then its expected value is and its variance is .
2. If are independent Normally distributed random variables each with the distribu-tion then:

because each of the variables has the standard Normal distribution .

1. Let be independent random variables each with the distribution and let , where at least one of the ’s is non-zero. Then the distribution of:

has larger mean and larger variance than where . This is called the **non-central chi-squared distribution** with degrees of freedom and **non-centrality parameter** . It is denoted by .

1. Suppose that the ’s are not necessarily independent and the vector has the multivariate normal distribution where the variance-covariance matrix is non-singular and its inverse is . Then:
2. More generally if then the random variable has the non-central chi-squared distribution where .
3. If are independent random variables with the chi-squared distributions , which may or may not be central, then their sum also has a chi-squared distribution with degrees of freedom and non-centrality parameter , i.e.,

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

1. Let , where has elements but the ’s are not independent so that is singular with rank and the inverse of is not uniquely defined. Let denote a generalized inverse of . Then the random variable has the non-central chi-squared distribution with degrees of freedom and non-centrality parameter .

**a.** As property 1 from above would suggest, the chi-squared distribution with degrees of freedom, , is the distribution of the sum of the squares of independent standard normal random variables. If  is a random variable following a normal distribution with mean and variance (), then the distribution of  is a special case of the **chi-squared distribution with one degree of freedom,** . Meaning that:

The chi-squared distribution with 1 degree of freedom is sometimes referred to as the exponential distribution with rate parameter (, ).

So, the distribution of is or equivalently, an exponential distribution with rate parameter .

**b.** The expression is the dot product of the vector with itself. So:

We know that and , and that they are independent. We also know (form a) that is a special case of the chi-squared distribution with one degree of freedom, , or in other words: .

Furthermore, we are given that: , thus:

Since both and 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 (and also according to property 3):

**c.** We know that and . Given that: and its distribution is , we have that:

The mean vector of , is:

While the Variance-Covariance matrix , is a diagonal matrix, because and are independent and it is:

Let us also compute the inverse of Variance-Covariance matrix , as it will be used:

Now, an expression for , will be:

As it was already shown above (in a), . Now, it was also shown (in b) that, and thus , plus a non-centrality parameter λ, which from property 6, is the following:

And therefore, since we are adding two chi-squared distributed variables, with one degree of freedom each, it follows that (again from property 6):

***Exercise 1.3:***

**Let the joint distribution of and be with:**

**a. Obtain an expression for . What is its distribution?**

**b.** **Obtain an expression for . What is its distribution?**

***SOLUTION:***

*A reminder from the book:*

1.4.2 Chi-squared distribution:

1. The **central chi-squared distribution** with degrees of freedom is defined as the sum of squares of independent random variables each with the standard Normal distri-bution. It is denoted by:

In matrix notation, if then so that .

1. If has the distribution , then its expected value is and its variance is .
2. If are independent Normally distributed random variables each with the distribu-tion then:

because each of the variables has the standard Normal distribution .

1. Let be independent random variables each with the distribution and let , where at least one of the ’s is non-zero. Then the distribution of:

has larger mean and larger variance than where . This is called the **non-central chi-squared distribution** with degrees of freedom and **non-centrality parameter** . It is denoted by .

1. Suppose that the ’s are not necessarily independent and the vector has the multivariate normal distribution where the variance-covariance matrix is non-singular and its inverse is . Then:
2. More generally if then the random variable has the non-central chi-squared distribution where .
3. If are independent random variables with the chi-squared distributions , which may or may not be central, then their sum also has a chi-squared distribution with degrees of freedom and non-centrality parameter , i.e.,

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

1. Let , where has elements but the ’s are not independent so that is singular with rank and the inverse of is not uniquely defined. Let denote a generalized inverse of . Then the random variable has the non-central chi-squared distribution with degrees of freedom and non-centrality parameter .

**a.** First, let us compute the inverse of Variance-Covariance matrix , as it will be needed. So:

Since and , then their difference shall be:

And therefore, the joint distribution, will have the following form:

From property 5, we know that for a multivariate normal distribution , the quadratic form follows a chi-squared distribution with degrees of freedom equal to the dimension of y (and in this case, we have only two dimensions), and therefore:

**b.** From the previous question (a), we already know that:

And therefore, the expression for shall be:

Now, the distribution of is a more general case of the one described in the previous question (a) and thus follows property 6, meaning that: “if then the random variable has the non-central chi-squared distribution where .”

In our case, can be written as , where: . So if we expanded on this, we would have:

with:

* , because , and the quadratic form of a multivariate normal distribution follows a chi-squared distribution with degrees of freedom equal to the dimension of (which is 2).
* is normally distributed with a mean of 0.

Thus is a sum of a chi-squared distribution and a normal distribution. This means that follows a non-central chi-squared distribution with 2 degrees of freedom and a non-centrality parameter , which is:

Therefore, in conclusion:

***Exercise 1.4:***

**Let be independent random variables each with the distribution . Let:**

**a. What is the distribution of ?**

**b. Show that .**

**c. From (b) it follows that . How does this allow you to deduce that and are independent?**

**d. What is the distribution of ?**

**e. What is the distribution of ?**

***SOLUTION:***

**a.** Since the​ are independent and each has the distribution , the expectation of is:

while its variance is:

We know that consists of a linear combination of independent, normally distributed variables and therefore it is itself normally distributed. Thus, the distribution of shall be:

**b.** Let us start from the definition of the sample variance:

**c.** We are given the following expression:

So, why are and independent? Let us look at the two right hand terms one by one.

Firstly, let us discuss the term:

Here is the sample variance, which is defined as:

Therefore:

The sample variance measures the spread of the individual ​’s around the sample mean . This involves degrees of freedom because the calculation of depends on data points, but the sample mean is used to estimate the center of the data, reducing the degrees of freedom by 1.

Thus, under the assumption that the ​​’s are normally distributed, the sum of squares, which was defined above, follows a chi-squared distribution with degrees of freedom:

Secondly, let us discuss the term:

And from question **a**, we already know that:

Therefore:

where is a standard normal random variable, . And hence:

Since the total sum of squares can be split into two independent components, one involving and the other involving , then by Cochran’s Theorem, the chi-squared terms must be independent. More formally, the independence of and implies that  **and are independent**.

*Cochran’s Theorem:*

Cochran's Theorem provides a way to decompose sums of squared normal random variables into independent chi-squared distributions. Specifically, if you have a set of independent normal random variables ​ drawn from . In our example, Cochran's Theorem states that the total sum of squares:

can be decomposed into two independent components:

This result is a key property of normal distributions and is a consequence of the fact that the sample mean and sample variance capture independent aspects of the data. captures location (center), while captures spread (variability) around the center.

**d.** As it was already shown in question **c**:

**e.** From question **c**, we got that:

Thus, the numerator follows a standard normal distribution, while the denominator involves the sample standard deviation, which is related to the chi-squared distribution with degrees of freedom. So, when we take the ratio of a standard normal random variable and the square root of a chi-squared random variable (divided by its degrees of freedom), the result follows a **t-distribution**. Hence:

***Exercise 1.5:***

**This exercise is a continuation of the example in Section 1.6.2 in which are independent Poisson random variables with the parameter .**

**a. Show that for .**

**b. Suppose** **. Find the maximum likelihood estimator of .**

**c. Minimize to obtain a least squares estimator of .**

***SOLUTION:***

**a.** We are given that  are independent Poisson random variables with the parameter . Therefore:

However, when , the whole term becomes zero, thus it is superfluous in our expression. We can take it out:

Setting , we get:

**b.** Given that ​ are independent Poisson random variables with parameter , the probability mass function for each ​ is:

The likelihood function is the product of the individual probabilities for all 's:

The log-likelihood function is the natural logarithm of the likelihood function:

To find the maximum likelihood estimator of , we take the derivative of with respect to and set it equal to zero:

**c.** To minimize , we need to take the derivative of with respect to and set it equal to zero, so in other words:

***Exercise 1.6:***

**This**

**a. Show that for .**

**b. Suppose . Find the maximum likelihood estimator of .**

**c. Minimize to obtain a least squares estimator of .**

***SOLUTION:***

**a.** We are