

5. Special Properties of Normal Samples

Suppose that $X = (X_1, X_2, \dots, X_n)$ is a **random sample** from the **normal distribution** with **mean** μ and **standard deviation** σ . In this section, we will establish several special properties of the **sample mean**, the **sample variance**, and some other important statistics. Since the sample (and in particular the sample size n) is fixed, it will be suppressed in the notation.

The Sample Mean

First recall that the sample mean is

$$M = \frac{1}{n} \sum_{i=1}^n X_i$$

The distribution of M follows easily from basic properties of independent normal variables:

1. Show that M is normally distributed with mean μ and variance $\frac{\sigma^2}{n}$.

The **standard score** of M is given as follows. This statistic will appear in several of the derivations in this section.

$$Z = \frac{M - \mu}{\sigma / \sqrt{n}}$$

2. Show that Z has the standard normal distribution.

The Sample Variance

The Distribution of the Special Sample Variance

Recall that if μ is known, a natural estimator of the variance σ^2 is the statistic

$$W^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

Although the assumption that μ is known is usually artificial, W^2 is very easy to analyze and it will be used in some of the derivations below.

3. Show that $\frac{n}{\sigma^2} W^2$ has the **chi-square distribution** with n degrees of freedom.

4. Use the result of the previous exercise to show that

- a. $\mathbb{E}(W^2) = \sigma^2$
- b. $\text{var}(W^2) = \frac{2\sigma^4}{n}$

Of course, these are special cases of the general results obtained in the section on [Sample Variance](#).

Independence of the Sample Mean and Variance

Recall that the standard version of the sample variance is the statistic

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - M)^2$$

The sample variance S^2 is the usual estimator of σ^2 when μ is unknown. The next series of exercises show that the sample mean M and the sample variance S^2 are [independent](#), a very important and useful property. First we will note a simple but interesting fact that holds for a random sample from any distribution, not just the normal.

5. Use basic [properties of covariance](#) to show that M and $X_i - M$ are uncorrelated for any $i \in \{1, 2, \dots, n\}$.

Our analysis hinges on the sample mean M and the vector of deviations from the sample mean. Let

$$Y = (X_1 - M, X_2 - M, \dots, X_{n-1} - M)$$

6. Show that S^2 can be written as a function of Y by noting that

$$\sum_{i=1}^n (X_i - M) = 0$$

7. Show that the M and the vector Y have a joint [multivariate normal distribution](#).

8. Use the results of the previous exercises to show that M and Y are independent.

9. Finally, show that M and S^2 are independent.

The Distribution of the Sample Variance

We can now determine the distribution of a simple multiple of the sample variance S^2 . Let

$$V = \frac{n-1}{\sigma^2} S^2$$

10. Show that $\frac{n}{\sigma^2} W^2 = V + Z^2$ where, as usual, Z is the standard score associated with the sample mean M and where W^2 is the special version of the sample variance. *Hint:* In the sum on the left, add and subtract M , and expand.

11. Show that V has the chi-square distribution with $n-1$ degrees of freedom. *Hint:* Use the result of the previous exercise, independence, and moment generating functions.

12. Use the result of the previous exercise to show that

a. $\mathbb{E}(S^2) = \sigma^2$

b. $\text{var}(S^2) = \frac{2\sigma^4}{n-1}$

Of course, these are special cases of the general results obtained in the section on [Sample Variance](#).

13. In the **basic random variable experiment**, select the chi-square distribution. Vary the degree of freedom parameter and note the shape and location of the probability density function and the mean, standard deviation bar. For selected values of the parameter, run the experiment 1000 times, updating every 10 runs, and note the apparent convergence of the empirical density and moments to the true distribution density and moments.

The T Statistic

The next sequence of exercises will derive the distribution of

$$T = \frac{M - \mu}{S / \sqrt{n}}$$

Note that T is similar to the standard score Z , but with the sample standard deviation S replacing the distribution standard deviation σ . The statistic T plays a critical role in constructing **interval estimates** for μ and performing **hypothesis tests** for μ .

14. As usual, let Z denote the standard score associated with the sample mean M and V the chi-square statistics associated with the sample variance S^2 . Show that

$$T = \frac{Z}{\sqrt{V/(n-1)}}$$

15. Use previous results to show that T has the **student t distribution** with $n - 1$ degrees of freedom.

16. In the **basic random variable experiment**, select the t distribution. Vary the degree of freedom parameter and note the shape and location of the probability density function and the mean, standard deviation bar. For selected values of the parameters, run the experiment 1000 times, updating every 10 runs, and note the apparent convergence of the empirical density and moments to the true distribution density and moments.

The F Statistic

The last statistic that we will study arises in the two-sample normal model. Thus, suppose that $\mathbf{X} = (X_1, X_2, \dots, X_m)$ is a random sample of size m from the normal distribution with mean μ and standard deviation σ and that

$\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ is a random sample of size n from the normal distribution with mean ν and standard deviation τ .

Finally, suppose that \mathbf{X} and \mathbf{Y} are independent. In the following exercises, we will use the basic notation established above, but we will indicate the dependence on the sample.

17. Show that the random variable given below has the **F distribution** with m degrees of freedom in the numerator and n degrees of freedom in the denominator:

$$F = \frac{W^2(\mathbf{X}) / \sigma^2}{W^2(\mathbf{Y}) / \tau^2}$$

18. Show that the random variable given below has the **F distribution** with $m - 1$ degrees of freedom in the numerator and $n - 1$ degrees of freedom in the denominator:

$$F = \frac{S^2(\mathbf{X}) / \sigma^2}{S^2(\mathbf{Y}) / \tau^2}$$

These variables are useful for [estimates](#) and [tests](#) of the ratio of the variances $\frac{\sigma^2}{\tau^2}$. The choice of the variables depends on whether the means μ and τ are known or unknown. Usually, of course, the means are unknown and so the statistic in [Exercise 18](#) is used.

19. In the [basic random variable experiment](#), select the F distribution. Vary the degrees of freedom parameters and note the shape and location of the probability density function and the mean, standard deviation bar. For selected values of the parameters, run the experiment 1000 times, updating every 10 runs, and note the apparent convergence of the empirical density and moments to the true distribution density and moments.

[Virtual Laboratories](#) > [5. Random Samples](#) > [1](#) [2](#) [3](#) [4](#) **[5](#)** [6](#) [7](#)

[Contents](#) | [Applets](#) | [Data Sets](#) | [Biographies](#) | [External Resources](#) | [Keywords](#) | [Feedback](#) | ©