THE DISTRIBUTIONS OF THE PRECISION CONSTANT AND ITS SQUARE IN SAMPLES OF n FROM A NORMAL POPULATION

By

H. M. Feldman
Washington University

INTRODUCTION

The following paper is a study of the properties of the distributions of the precision constant and its square in samples of from a normal population. The properties studied are (1) modes and optimum values, (2) the first four moments, (3) skewness and flatness, and (4) medians and quartiles.

The distribution curves shown in the figure are for $n=4$, 10, and 25. All the curves are drawn together and to the same scale, so that a graphical comparison of the two distributions can be easily made for both the same and different values of n. The numerical values for the various parameters given in the tables are for $n=4$, 10, 25, and 100, except in the case of the medians and quartiles where the values for $n=100$ are omitted, and in case of $n=4$, no moments higher than the second exist for the precision constant, and none higher than the first for the precision constant squared distribution.

1. *Distributions*

Let us denote the standard deviation, precision constant, and the precision constant squared of the parent population by S, H, and U, respectively, and those of a sample from the given population by s, h, and u, respectively. The standard deviation, s, is then defined in terms of the variates, $x_1, x_2, \cdots, x_n$ and its mean $\bar{x}$ by the equation

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (1.0)$$

We also have the following well known relations between S, H, and U, or s, h, and u:

$$s = \frac{1}{\sqrt{2}\,h} \quad \ldots\ldots\ldots\ldots \quad (1.1)$$

$$s^2 = \frac{1}{2u}, \text{ or } u = h^2 \quad \ldots\ldots \quad (1.2)$$

The distribution of s as given by R. A. Fisher* and others is

$$f(s)ds = \frac{n^{\frac{n-1}{2}}\, s^{n-2}\, e^{-\frac{nS^2}{2s^2}}\, ds}{2^{\frac{n-3}{2}}\, S^{n-1}\, \Gamma\left(\frac{n-1}{2}\right)} \quad \ldots \quad (1.3)$$

where n is the number of items in the sample.

The distribution of s and the transformations (1.1) and (1.2) enable us to find the distributions of the precision constant h, and its square u. Thus, using (1.1) and (1.3) we get

$$F(h)dh = \frac{n^{\frac{n-1}{2}}\, H^{n-1}\, h^{-n}\, e^{-\frac{nH^2}{2h^2}}\, dh}{2^{\frac{n-3}{2}}\, \Gamma\left(\frac{n-1}{2}\right)} \quad \ldots \quad (1.4)$$

and by means of (1.2) and (1.3) we find

$$\phi(u)du = \frac{n^{\frac{n-1}{2}}\, U^{\frac{n-1}{2}}\, u^{-\frac{(n+1)}{2}}\, e^{-\frac{nU}{2u}}\, du}{2^{\frac{n-1}{2}}\, \Gamma\left(\frac{n-1}{2}\right)} \quad \ldots \quad (1.5)$$

2. *Modal and Optimum Values*

We shall now obtain some of the properties of the distributions of our parameters. The simplest of the properties of any

*See, for example, R. A. Fisher, Applications of "Student's" distribution, Metron, Vol. 5, No. 3, (Dec. 1, 1925), pp. 90-104.

continuous distribution is its modal value or the abscissa of the maximum ordinate of the curve. Denoting the modal values of the distributions (1.4) and (1.5) by $\tilde{h}$ and $\tilde{u}$ respectively, we find from the condition for an extremum

$$\frac{dF}{dh}=0, \quad \frac{d\phi}{du}=0,$$

$$H=\tilde{h} \quad \ldots\ldots\ldots\ldots\ldots \quad (2.0)$$

$$U=\frac{n+1}{n}\tilde{u} \quad \ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

In obtaining $\tilde{h}$ and $\tilde{u}$ we regard h and $\bar{u}$ as variables and H and U as constants. We may, however, reverse our point of view. That is, we regard H and U as variables and h and u as constants. In that case the right hand sides of (1.4) and (1.5) become functions of H and U, and from

$\frac{dF}{dH}=0$, and $\frac{d\phi}{dU}=0$, we get

$$h=\sqrt{\frac{n}{n-1}}\,\hat{H} \quad \ldots\ldots\ldots\ldots\ldots \quad (2.2)$$

$$u=\frac{n}{n-1}\hat{U} \quad \ldots\ldots\ldots\ldots\ldots \quad (2.3)$$

The quantities $\hat{H}$ and $\hat{U}$ R. A. Fisher calls the ***optimum*** values.

3. *Moments*

Precision Constant

In order to distinguish between the moments of the two distributions treated in this paper, we shall denote the κth moment of the precision constant about the origin by $\mu_i'(n)$ and that of its square by $\mu_i'(u)$ with similar notation for the moments about the mean.

Using the general definition of a moment of a continuous distribution, we obtain for the first moment of n, which is also its mean

$$\mu_1'(h) = \frac{n^{\frac{n-1}{2}} H^{n-1}}{2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right)} \int_0^\infty h^{-n} e^{-\frac{nH^2}{2h^2}} dh \quad \ldots\ldots (3.10)$$

To put this into an integrable form we make the transformation

$$t = \frac{nH^2}{2h^2} \quad \ldots\ldots\ldots\ldots\ldots (3.11)$$

This yields for the first moment

$$u_1'(h) = H\sqrt{\frac{n}{2}} \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \quad \ldots\ldots\ldots\ldots (3.12)$$

To facilitate calculation we express this in terms of factorials. For this purpose we have two cases to consider, namely, the case when n is even, and that when n is odd.

When n is even $\frac{n-2}{2}$ is an integer, and,

$$\mu_1'(h) = H\sqrt{\frac{n}{2}} \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} = H\sqrt{n} \frac{(n-4)(n-6)\cdots 2}{(n-3)(n-5)\cdots 1} \sqrt{\frac{2}{\pi}}$$

$$= H\sqrt{n} \frac{\left[2^{\frac{n-4}{2}} \left(\frac{n-4}{2}\right)!\right]^2}{(n-3)!} \sqrt{\frac{2}{\pi}} \quad \ldots\ldots (3.13)$$

When n is odd $\frac{n-1}{2}$ is an integer and hence

$$\mu_1'(h) = H\sqrt{\frac{n}{2}} \frac{\Gamma \frac{n-2}{2}}{\Gamma \frac{n-1}{2}} = H\sqrt{n} \frac{(n-4)(n-6)\cdots 1}{(n-3)(n-5)\cdots 2} \sqrt{\frac{\pi}{2}}$$

$$= H\sqrt{n}\, \frac{(n-3)}{\left[2^{\frac{n-3}{2}}\ \frac{n-3}{2}!\right]}\, 2\sqrt{\frac{\pi}{2}} \qquad \ldots\ldots (3.14)$$

Similarly we obtain for the second, third, and forth moments of the distribution of about the origin the following expressions:

$$\mu_2'(h) = H^2\,\frac{n}{2}\,\frac{\Gamma\left(\frac{n-3}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} = \frac{n}{n-3}\,H^2, \qquad \ldots\ldots\ldots\ldots (3.15)$$

$$\mu_3'(h) = H^3\left(\frac{n}{2}\right)^{3/2}\frac{\Gamma\left(\frac{n-4}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} = \frac{n}{n-4}\,\mu_1'(h)\,H^3, \qquad \ldots\ldots (3.16)$$

$$\mu_4'(h) = H^4\left(\frac{n}{2}\right)^2\frac{\Gamma\left(\frac{n-5}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} = \frac{n^2}{(n-3)(n-5)}\,H^4 \qquad \ldots\ldots (3.17)$$

Moments about the Mean

To study such properties of a distribution curve as skewness and flatness we must have the moments of the curve about the mean. To obtain these we use the well known formulae for expressing the moments about the mean in terms of the moments about any origin. Using these formulae we obtain for the first four moments of the precision constant, h, about the mean the following expressions:

$$\mu_1'(h) = 0$$

$$\mu_2(h) = \frac{\left[n-(n-3)\mu_1'^2(h)\right]H^2}{(n-3)} \qquad \ldots\ldots\ldots\ldots (3.18)$$

$$\mu_3(h) = \frac{\left[2(n-3)(n-4)\mu_1'^2 - n(2n-9)\mu_1'(n)\right]H^3}{(n-3)(n-4)} \qquad \ldots\ldots (3.19)$$

$$\mu_4(h) = \frac{[n^2(n-4)+2n(n-6)(n-5)\mu_1'^2 - 3(n-3)(n-4)(n-5)\mu'^4]H^4}{(n-3)(n-4)(n-5)} \qquad \ldots . (3.20)$$

where $\mu_1'(h)$ is given by (3.12).

For future use we shall give here approximations for $\mu_1'(h)$, $\mu_2(h)$, $\mu_3(h)$, and $\mu_4(h)$. The approximations were obtained by expanding the various quantities into power series of $\frac{1}{n}$. The derivation of these are not difficult but rather long and will therefore not be given in this paper.

These approximations are as follows:

$$\mu_1'(h) = \left(1 + \frac{5}{4n} + \frac{49}{16n^2}\right) H \qquad \ldots . (3.21)$$

$$\mu_2(h) = \frac{1}{2n} H^2 \qquad \ldots . (3.22)$$

$$\mu_3(h) = \frac{a}{n^2} H^3 \qquad \ldots . (3.23)$$

(where a is a constant)

$$\mu_4(h) = \frac{3}{4n^2} H^4 \qquad \ldots . (3.24)$$

Precision Constant Squared

The first moment of the precision constant squared distribution is, using the general definition of a continuous distribution curve about the origin

$$\mu_1'(u) = \frac{n^{\frac{n-1}{2}} U^{\frac{n-1}{2}}}{2^{\frac{n-3}{2}} \Gamma \frac{n-1}{2}} \int_0^\infty u^{-\frac{n+1}{2}} e^{-\frac{nU}{2u}} du \qquad \ldots . (3.30)$$

We reduce this to a known integral by means of the transformation

$$t = \frac{nU}{2u} \qquad \ldots . (3.31)$$

We then obtain for the first four moments of u about the origin the following simple expressions:

$$\mu_1'(u) = \frac{n}{n-3} U \quad \ldots \quad (3.32)$$

$$\mu_2'(u) = \frac{n^2}{(n-3)(n-5)} U^2 \quad \ldots \quad (3.33)$$

$$\mu_3'(u) = \frac{n^3}{(n-3)(n-5)(n-7)} U^3 \quad \ldots \quad (3.34)$$

$$\mu_4'(u) = \frac{n^4}{(n-3)(n-5)(n-7)(n-9)} U^4 \quad \ldots \quad (3.35)$$

Moments about the Mean

For the moments about the mean of the precision constant squared distribution we have:

$$\mu_1 = 0$$

$$\mu_2 = \frac{2n^2}{(n-3)^2(n-5)} U \quad \ldots \quad (3.36)$$

$$\mu_3 = \frac{16n^3}{(n-3)^3(n-5)(n-7)} U^3 \quad \ldots \quad (3.37)$$

$$\mu_4 = \frac{4n^4(2n+27)}{(n-3)^4(n-5)(n-7)(n-9)} U^4 \quad \ldots \quad (3.38)$$

Skewness and Flatness

From the above expressions for the mean and also from the numerical values given in the tables we may conclude that the precision constant distribution is less skew than the distribution of the precision constant squared, at least for values of n up to 100. But what happens when n grows very large? To answer this we make use of the Pearsonian measure of skewness, β_1, defined by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \quad \ldots \quad (3.40)$$

Since we are now interested in large values of n we make use of the approximate values of $\mu_2(h)$, and $\mu_3(h)$, which are

$$\mu_2(h) = \frac{1}{2n} H^2, \text{ and } \mu_3(h) = \frac{a}{n^2}$$

Hence we get

$$\lim_{n \doteq \infty} \beta_1(h) = \lim_{n \doteq \infty} \frac{a^2 H^6}{n^4} \div \frac{1}{n^3} H^6 = \lim_{n \doteq \infty} \frac{a^2}{n}$$

and $\displaystyle \lim_{n \doteq \infty} \frac{a^2}{n} = 0$ . . . . (3.41)

To find $\lim_{n \doteq \infty} \beta_1(u)$ we make use of the exact values of $\mu_2(u)$ and $\mu_3(u)$ which are given by (3.36) and (3.37). This gives

$$\lim_{n \doteq \infty} \beta_1(u) = \lim_{n \doteq \infty} \left[\frac{16 n^3 U^3}{(n-3)^3(n-5)(n-7)}\right]^2 \div \left[\frac{2n^2 U^2}{(n-3)^2(n-5)}\right]^3$$

$$= \lim_{n \doteq \infty} \frac{32(n-5)}{(n-7)^2} = 0 \quad \text{. . . . . (3.42)}$$

From (3.41) and (3.42) we learn that both the precision constant distribution, and that of its square, approach perfect symmetry as the size of the sample, n, approaches infinity.

The flatness or kurtosis of a curve is measured by the quantity

$$\beta_2 = \frac{\mu_4}{\mu_1^2} \quad \text{. . . . . (3.43)}$$

From this and the expressions (3.22 and (3.24), (3.33) and (3.35) we get

$$\lim_{n \doteq \infty} \beta_2(h) = \lim_{n \doteq \infty} \frac{3H^4}{4n^2} \div \left(\frac{H^2}{2n}\right)^2 = 3 \quad \text{. . . . . (3.44)}$$

and

$$\lim_{n \doteq \infty} \beta_2(u) = \lim_{n \doteq \infty} \frac{4n^4(2n+27)U^4}{(n-3)^4(n-5)(n-7)(n-9)} \div \frac{4n^4 U}{(n-3)^4(n-5)}$$

$$= \lim_{n \doteq \infty} \frac{(n-5)(2n+27)}{(n-7)(n-9)} = 2 \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (3.45)$$

We may conclude, then, that while the distributions of the precision constant and its square are both perfectly symmetrical for very large values of n, they are nevertheless entirely distinct distribution curves for both small and large values of n, since $\lim\limits_{n \doteq \infty} \beta_2(h) = 3$, and $\lim\limits_{n \doteq \infty} \beta_2(u) = 2$. As a matter of fact the distribution of the precision constant approaches the normal curve, while the precision constant squared distribution approaches a curve of the form

$$y = y_0 \left(1 - \frac{x^2}{a^2}\right)^{\frac{1}{2}}$$

where y_0 and a are constants.

4. *Quartiles and Medians*

The quartiles of a continuous distribution $f(x)$ may be defined by the equation

$$\int_0^{Q_i} f(x)\,dx = \frac{i}{4}, \quad (i = 1, 2, 3) \quad . \; . \; . \; . \quad (4.00)$$

For $i = 1$, Q_i is called the lower quartile; for $i = 2$, Q_i is called the median and for $i = 3$, Q_i is called the upper quartile.

In order to find the quartiles of the distributions studied in this paper we must make use of the incomplete Γ- function. This function is defined as follows

$$I(u,p) = \frac{1}{\Gamma(p+1)} \int_0^{u\sqrt{1+p}} e^{-v} v^p\,dv \quad . \; . \; . \; . \quad (4.10)$$

Pearson's "Tables of the Inconplete Γ-Function" give the values of I, u, and p. Thus, if we know any two of the variables we can easily find the third.

Let us take, now, the distribution of h,

$$F(h)dh = \frac{n^{\frac{n-1}{2}} H^{n-1} h^{-n} e^{-\frac{nH^2}{2h^2}} dh}{2^{\frac{n-3}{2}} \Gamma\left(\frac{n-1}{2}\right)}$$

The various quartiles of this distribution will be given by

$$n^{\frac{n-1}{2}} H^{n-1} \int_0^{Q_i} h^{-n} e^{-\frac{nH^2}{2h^2}} dh = \frac{i}{4} \quad \ldots . \ (4.11)$$

By making the transformation

$$v = \frac{nH^2}{2n^2}$$

(4.11) is reduced to

$$\frac{1}{\Gamma\left(\frac{n-1}{2}\right)} \int_{\infty}^{u_i\sqrt{\frac{n-1}{2}}} e^{-v} v^{\frac{n-3}{2}} = \frac{i}{4} \quad \ldots \ldots . \ (4.13)$$

This is, of course, $I(u_i, \frac{u-3}{2})$, and since i and u are known we can easily find u_i from Pearson's Tables.

Comparing (4.11) and (4.13) and taking into account the transformation (4.12) we find that

$$Q_i = \frac{n^{\frac{1}{2}}}{u_i^{\frac{1}{2}} (2n-2)^{\frac{1}{4}}} H.$$

Since the lower limit in (4.13) is ∞ instead of O we see that the lower and upper quartiles are reversed.

To find the quartiles for the distribution of u we simply make use of the relation $u = h^2$.

In conclusion we may state that the results of this paper are similar to those of Professor Rietz on the distributions of the standard deviation and the variance.*

---

*Rietz, H. L. A comparison of the distributions curves of variance and of standard deviation, *Mathematical Monthly,* Vol. 36 (August-Sept., 1929), p. 355.

TABLE I

Modal and optimum values for the distribution of the precision constant and its square from a normal population for samples of 4, 10, 25, and 100.

	Pr. const. h		Pr. const. squared u	
n	Mode	Optimum	Mode	Optimum
4	H	$1.154H$	$0.800\,U$	$1.333\,U$
10	H	$1.054\,H$	$0.909\,U$	$1.111\,U$
25	H	$1.021\,H$	$0.962\,U$	$1.042\,U$
100	H	$1.005\,H$	$0.990\,U$	$1.010\,U$

TABLE II

Values of the mean and the first four moments about the mean of the distributions of the precision constant and its square for samples of 4, 10, 25, and 100 from a normal population.

Precision constant h				
n	Mean μ_1'	μ_2	μ_3	μ_4
4	$1.596\,H$	$1.454\,H^2$		
10	$1.153\,H$	$0.0982\,H^2$	$0.0678\,H^3$	$0.0815\,H^4$
25	$1.054\,H$	$0.0255\,H^2$	$0.0031\,H^3$	$0.0026\,H^4$
100	$1.013\,H$	$0.0053\,H^2$	$0.000141H^3$	$0.0000907\,H^4$

Precision constant square U				
n	μ_1'	μ_2	μ_3	μ_4
4	$4.000\,U$			
10	$1.429\,U$	$0.816\,U^2$	$3.110\,U^3$	$52.200\,U^4$
25	$1.136\,U$	$0.129\,U^2$	$0.06522U^3$	$0.892\,U^4$
100	$1.031\,U$	$0.224\,U^2$	$0.0199\,U^3$	$0.00129U^4$

TABLE III

Values of β_1 and β_2 for the distribution of the precision constant and its square for samples of 10, 25, 100, and ∞ from a normal population.

	h		u	
n	β_1	β_2	β_1	β_2
10	4.8548	8.4521	17.7807	78.3416
25	0.7596	3.9982	2.0099	5.3172
100	0.00134	3.2289	0.3515	2.5476
∞	0.0000	3.0000	0.0000	2.0000

TABLE IV

Medians and quartiles for the distributions of the precision constant and its square for samples of 4, 10, and 25 from a normal population.

	Precision constant h			Precision constant squared u		
n	Q_1	Median	Q_3	Q_1	Median	Q_3
4	0.985 H	1.298 H	1.816 H	0.970 U	1.685 U	3.298 U
10	0.937 H	1.094 H	1.230 H	0.878 U	1.197 U	1.513 U
25	0.941 H	1.036 H	1.146 H	0.885 U	1.073 U	1.313 U

Distribution curves of the precision constant and its square for samples of 4, 10, and 25, from a normal population.

Note that the mode for the precision constant distribution is the same for all values of n as is seen from (2.0).

The solid curves are for the precision constant, the dotted curves for its square.

The unit used is the standard deviation of the population