MINIMAL VARIANCE AND ITS RELATION TO EFFICIENT MOMENT TESTS

By J. R. Vatnsdal

State College of Washington

1. Summary. When a curve is fitted to a set of data by moments, the usual procedure used in testing the hypothesis that the population is of the given form with the parameters as computed from the moments is to compare the higher moments with their expected values as determined by the hypothesis. Generally speaking, moments about the mean are computed although the reason for this is not clear. To shed some light on this question, the sample given in the introduction is fitted to two curves. Moments about various points are compared with their expected values and the discrepancy in standard units examined. This discrepancy is found to vary widely and to have a maximum. The notion of equivalent moment tests is introduced, and on this basis the most efficient moment test is defined in such a way that of all equivalent moment tests, this one is most likely to reject a false hypothesis.

For any moment it is shown that there is a point about which its variance is a minimum. The conditions are found which determine the position of this point for second and third moments. It is proved that for symmetrical populations the variance is minimal when the moments are computed about the mean of the population. If the population is an asymmetrical Pearson frequency function, it is proved that the point about which the third moment variance is minimal differs more from the mean than does the corresponding point for second moments. The condition is pointed out for which this is true in the general case.

The third and fourth standard semi-invariants of second moments of minimal variance are computed and compared to those of the second moment about the mean. The ratios of these are displayed for some populations to illustrate how this may be used to investigate when the approach to normality is more rapid in one case than in the other. Some examples are presented to contrast these and other tests.

2. Introduction. In testing the hypothesis that a given set of observations is a random sample from a completely specified population (either a priori or specified by a consideration of the sample), generally the Chi-square test is applied or certain functions of the moments are compared with their expected values and the significance of their departure as determined by the hypothesis is examined.

In the Neyman-Pearson theory it is required that the functional form be known. The hypothesis then is some statement concerning the parameters. The main principle there used is that the test used should be such that, while keeping the probability of rejecting the hypothesis when true at a certain significance level, it will minimize the chance of accepting the hypothesis when some alternative is true.

However, if the functional form is regarded as unknown, the alternative hypotheses are then usually unknown. The test then must be one that does not depend on alternatives. In the light of incomplete knowledge of the distribution of sample statistics, and since moments of moments are practically the only ones known, we shall here use the principle of comparing observed moments with their expected values. It is known that the distribution of moments in large samples is asymptotic to the normal distribution if the appropriate moments of the population exist [1]. Here we shall confine ourselves to such populations and large samples.

To introduce the idea which underlies the theory here presented, consider a simple example. Suppose a sample is given and the hypothesis is of the form $f(x, \theta)$ with $\theta = \theta_0$. Furthermore, suppose the first moment of the sample is equal to its expected value. If a second-moment test is used, this means that one computes the arithmetic mean of the squares of the deviations of the elements of the sample about some point, and compares this with the theoretical moment about the same point. Generally speaking, the point used is the mean of the population or the mean of the sample. However, the point may be chosen in any manner. For each such choice a test can be devised such that the probability of rejecting the hypothesis when true is ϵ. All such tests are called equivalent moment tests. Among these equivalent moment tests, one particular second-moment will have the minimal variance. This one is here called the most efficient moment test.

This test has the property that the range of values of the second moment for which the hypothesis is accepted is as small as possible. Thus of all equivalent second-moment tests, this one is most likely to reject a false hypothesis.

This idea may be easily extended to moments of higher order, in all of which the concept of minimal variance is fundamental. The point of view may be taken that the point about which the moments are computed should be such that the variance is a minimum, or what is equivalent, the variance of moments about the origin is minimized by choosing the origin properly.

An example is here presented to bring this out more clearly. A sample of 1,000 items is given and fitted by the first two moments to two different frequency functions. (The sample items are not given here; they are to be found in *Tables for Statisticians* [2]). The third and fourth moments have been computed and the discrepancies in standard units as determined by the hypotheses are exhibited in a table.

This sample of 1,000 items considered as a sample from an infinite population has these moments:

$$m_1' = 139.288$$
$$m_2' = 19692.452$$
$$m_3' = 2827467.388$$
$$m_4' = 412561061.04$$

By fitting the first two moments of the sample to curve A,

$$y = \frac{a^{n+1}}{\Gamma(n+1)} x^n e^{-ax}$$

we get $a = 0.4781516735$ and $n = 65.60079029$; to curve B,

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

we get $\mu = 139.288$ and $\sigma^2 = 291.305056$.

The discrepancy between the observed and theoretical rth moment about any point is measured by

$$t = \frac{m_r'' - \mu_r''}{\sqrt{\dfrac{\mu_{2r}'' - \mu_r''^2}{n}}}$$

in which m_r'' is the rth moment of the sample of n about this point, and μ_r'' is the rth moment of the population about the same point.

The values of $| t |$ have been computed corresponding to various points for the third and fourth moments. These are exhibited in four tables, given below.

Examination of the table for the discrepancy between the observed and theoretical third moments for curve B, shows that when this moment is computed about $x = 0$, the hypothesis is accepted at the 1% level; this is also true for $x = 39.3$, but for $x = 139.3$ the hypothesis would be rejected at that level. It is evident that some rule must be established to decide what point is to be used to make the test.

If the curve is fitted by the first two moments the value $m_3'' - \mu_3''$ is the same for every point. This is easily demonstrated, for if m_3'' and μ_3'' are measured about a point h units to the right of the origin, $m_3'' = m_3' - 3hm_2' + 3h^2m_1' - h^3$ and $\mu_3'' = \mu_3' - 3h\mu_2' + 3h^2\mu_1' - h^3$. Now, $m_2' = \mu_2'$ and $m_1' = \mu_1'$. It follows that $m_3'' - \mu_3'' = m_3' - \mu_3'$.

The maximum value of $| t |$ is attained when the variance of third moments is a minimum. In this manner it is assured that the range of values for which the third moment is accepted shall be a minimum.

If the third moments agree, or the agreement is sufficiently close such that the hypothesis cannot be rejected, $m_4'' - \mu_4''$ is constant or varies only slightly from point to point, so that minimizing the variance yields the maximum value of t.

As is seen from the tables above, when the moments are compared at the different points, the hypothesis may be accepted for one point and rejected for another. By the principle of using the point which yields the minimal variance, the hypothesis will be rejected more often than for other points. Thus, of all equivalent moment tests, this one is most likely to reject a false hypothesis.

The problem of determining for various moments how the origin may be chosen such that the variance of the distribution of these moments shall be a minimum is now considered.

3. First moments. In the case of the first moment, whose expected value is the mean of the population, the variance is given by $\frac{1}{n}(\mu_2' - \mu_1'^2)$. It is obvious that the choice of origin does not affect the variance of the first moment, since it is well known that $\mu_2' - \mu_1'^2$ is invariant with respect to choice of origin.

4. Minimal variance of second moments. The variance of second moments about an arbitrary origin is $\frac{1}{n}(\mu_4' - \mu_2'^2)$. Expressed in terms of μ_1' and central

TABLES

Curve A.

Third moments.		Fourth moments.	
Point	t	Point	t
0	.0365	0	.197
50	.084	50	.697
100	.33	100	4.74
120	.77	120	14.17
130	1.28	130	26.76
140	1.91	140	49.03
142	1.95	145	45.26
145	1.90	150	42.89
150	1.60	160	21.31
160	.95	180	6.25
170	.57	200	2.51
180	.37	300	.183
200	.18		

Curve B.

Third moments.		Fourth moments.	
Point	t	Point	t
0	.085	0	.02
39.3	.19	39.3	.13
89.3	.69	99.3	.88
109.3	1.16	109.3	1.09
119.3	2.39	119.3	2.00
129.3	4.05	129.3	3.18
139.3	5.57	133.3	3.83
149.3	4.05	135.3	3.96
159.3	2.39	137.3	3.93
169.3	1.16	139.3	3.67
179.3	.98	140.3	3.46
189.3	.69	143.3	2.72
199.3	.50	148.3	1.59
209.3	.38	159.3	.39
239.3	.19	179.3	.13
		239.3	.07

moments, this may be written

(1) $$\mu_2(m_2') = \frac{1}{n}(\mu_4 - \mu_2^2 + 4\mu_3\mu_1' + 4\mu_2\mu_1'^2).$$

Here it is evident that the variance of second moments does depend on the choice of origin, and is not invariant under translation.

The minimum value of $\mu_2(m_2')$ is given by $\mu_1' = -\frac{\mu_3}{2\mu_2}$ and is $\frac{1}{n}\left(\mu_4 - \mu_2^2 - \frac{\mu_3^2}{\mu_2}\right)$. Then we may write

(2) $$\mu_2(m_2^*) = \frac{1}{n}\left(\mu_4 - \mu_2^2 - \frac{\mu_3^2}{\mu_2}\right).$$

Throughout this paper m_2^* denotes the second moment of the sample about an origin chosen such that $\mu_1' = -\frac{\mu_3}{2\mu_2}$, which is the value of μ_1' which minimizes (1); m_2^0 denotes the second moment about an origin chosen such that $\mu_1' = 0$; m_2 denotes the second moment about the mean of the sample. It may be noted that in large samples the distributions of m_2^0 and m_2 are approximately the same.

It is clear from (2) that if $\mu_3 = 0$, or, if the population is symmetric, i.e. $f(-x) = f(x)$, then $\mu_2(m_2^*) = \mu_2(m_2^0)$. However, if $\mu_3 \neq 0$ then $\mu_2(m_2^*) < \mu_2(m_2^0)$.

5. A moment inequality. Since the quantity given by (2) is essentially non-negative, an inequality is obtained valid for any distribution in which the first four moments exist, viz.

$$\mu_4 - \mu_2^2 - \frac{\mu_3^2}{\mu_2} \geq 0, \qquad \mu_2 \neq 0 \tag{3}$$

or in standard moments

$$\alpha_4 - \alpha_3^2 - 1 \geq 0. \tag{4}$$

This is a stronger inequality than the one given by Bertelsen [3], i.e. $\alpha_3^2 - \alpha_4 - 2 < 0$ or the one generally known, $\alpha_4 \geq \alpha_3^2$, [4]. This inequality, however, was known to K. Pearson [5, p. 432], although he derived it from a different point of view.

6. Minimal variance of higher moments. The variance of the distribution of rth moments of random samples about an arbitrary origin always has a minimum. The variance of m_r' is given by

$$\mu_2(m_r') = \frac{1}{n}(\mu_{2r}' - \mu_r'^2). \tag{5}$$

This expression when expanded in powers of μ_1' is always a polynomial of even degree with the coefficient of the highest power a positive number. Furthermore, by differentiating $\mu_2(m_r')$ with respect to μ_1' and equating the derivative to zero, the value of μ_1' which minimizes $\mu_2(m_r')$ will be found among the solutions of that equation.

For third moments of samples the variance is given by

$$\mu_2(m_3') = \frac{1}{n}[\mu_6' - \mu_3'^2]$$

which, when expressed in terms of moments about the mean and powers of the mean, becomes

$$\mu_2(m_3') = \frac{1}{n}[\mu_6 - \mu_3^2 + 6(\mu_5 - \mu_3\mu_2)\mu_1' + (15\mu_4 - 9\mu_2^2)\mu_1'^2 + 18\mu_3\mu_1'^3 + 9\mu_2\mu_1'^4]. \tag{6}$$

Differentiating with respect to μ_1' and equating to zero, we have

$$6\mu_2\mu_1'^3 + 9\mu_3\mu_1'^2 + (5\mu_4 - 3\mu_2^2)\mu_1' + (\mu_5 - \mu_3\mu_2) = 0. \tag{7}$$

By straightforward application of the methods of solving cubics, it is easy to show by means of (3) that (7) has one real root only, which moreover is $\lesseqgtr -\dfrac{\mu_3}{2\mu_2}$ according as

$$\alpha_5 - \alpha_3(\tfrac{5}{2}\alpha_4 - \tfrac{3}{2}\alpha_3^2 - 1) \gtreqless 0.$$

Since it can also be shown by means of (3) that the second derivative of (6) is positive, this root of (7) will minimize $\mu_2(m_3')$.

These facts demonstrate:

THEOREM I. *The point about which the arithmetic mean of the cubes of the variates has minimal variance is to the right, at, or to the left of the corresponding point for the squares according as*

$$\alpha_5 - \alpha_3(\tfrac{5}{2}\alpha_4 - \tfrac{3}{2}\alpha_3^2 - 1) \lesseqgtr 0. \tag{8}$$

By examination of (7) it is readily seen that if $\alpha_5 = \alpha_3$ or if the population is symmetric, the real root will be zero; so that for such a population the variance of third moments is a minimum when moments are taken about the mean of the population. If $\alpha_5 \neq \alpha_3$ the variance of third moments will be a minimum when taken about some other point.

For fourth moments of samples the variance is of the sixth degree in μ_1' and its derivative therefore of the fifth degree. There is not much to be said in a general way except that if $\alpha_7 = \alpha_4\alpha_3$ or if the population is symmetric, $\mu_1' = 0$ will cause this derivative to vanish.

If the distribution is a Pearson frequency function, from the recursion formula for the moments [6, p. 24],

$$\alpha_5 = \alpha_3\left(\frac{2\alpha_4 + 4 + 2\delta}{1 - \delta}\right)$$

where

$$\delta = \frac{2\alpha_4 - 3\alpha_3^2 - 6}{\alpha_4 + 3}.$$

The criterion (8) can be written

$$\alpha_3\left(\frac{2\alpha_4 + 4 + 2\delta}{1 - \delta}\right) + \alpha_3 + \tfrac{3}{2}\alpha_3^3 - \tfrac{5}{2}\alpha_4\alpha_3. \tag{9}$$

It will now be shown that (9) $\gtreqless 0$ according as $\alpha_3 \gtreqless 0$, since (9) is $\alpha_3 D$ where

$$D = \frac{2\alpha_4 + 4 + 2\delta}{1 - \delta} + 1 + \tfrac{3}{2}\alpha_3^2 - \tfrac{5}{2}\alpha_4. \tag{10}$$

It suffices to show that $D > 0$ for all Pearson curves. Using the method of Lagrange multipliers, it is possible to show that within the permissible range of values of the variables involved, the g.l.b. of D is $\frac{1}{2}$, and so $D > 0$. It has been proved that the variance of the squares is a minimum when $\mu_1' = \frac{-\alpha_3}{2}\sigma$. It has just been shown that the sign of (9) agrees with that of α_3. These, together with Theorem I, demonstrate

THEOREM II. *For Pearson frequency functions, $\alpha_3 \neq 0$, the point about which the variance of cubes is a minimum deviates more from the mean than does the corresponding point for the squares.*

7. Symmetric populations. For the distribution of rth moments of samples

$$\mu_2(m_r') = \frac{1}{n}(\mu_{2r}' - \mu_r'^2). \tag{11}$$

To find the minimum of (11) expand in terms of central moments and powers of μ_1', differentiate with respect to μ_1', and equate to zero. This yields:

$$\begin{aligned}(2r-2)r^2\mu_2\mu_1'^{2r-3} + \cdots + K\mu_1'^{K-1}\Bigg[\binom{2r}{K}\mu_{2r-K} \\ - \sum_{i=0}^{K}\binom{r}{i}\binom{r}{K-i}\mu_{r-i}\mu_{r-K+i}\Bigg] + \cdots + 2r(\mu_{2r-1} - \mu_r\mu_{r-1}) = 0.\end{aligned} \tag{12}$$

For each power of μ_1', the coefficient is an isobaric moment function and is of even weight when the power of μ_1' is odd, and of odd weight when the power of μ_1' is even. If the population is symmetric the coefficients of even powers will vanish as will the constant term. Then μ_1' will be a factor, the other factor being a polynomial with only even powers of μ_1'. In this latter factor, where K is even, the coefficient of $K\mu_1'^{K-2}$ is

$$\binom{2r}{K}\mu_{2r-K} - \sum_{i=0}^{K}\binom{r}{i}\binom{r}{K-i}\mu_{r-i}\mu_{r-K+i}\,. \tag{13}$$

Since

$$\binom{x+y}{n} = \sum_{m=0}^{n}\binom{x}{n-m}\binom{y}{m},$$

(13) may be written

$$a\mu_{2r-K} + \sum_{i=0}^{K} b_i(\mu_{2r-K} - \mu_{r-i}\mu_{r-K+i}), \qquad r-i,\ K \text{ even},$$

where a, b_i are non-negative integers.

It can be immediately established by use of an inequality due to Tchebycheff [7, pp. 43, 168] that $\mu_{2K+2l}' \geq \mu_{2K}'\cdot\mu_{2l}'$ and therefore (13) is positive or zero.

To sum up, if the odd moments vanish (12) will have a factor μ_1' and a factor which is a polynomial with even powers only of μ_1' with positive coefficients; therefore there is one and only one solution, $\mu_1' = 0$. This establishes

THEOREM III. *For a symmetrical population, the distribution of* rth *moments of samples has minimal variance when the origin is the population mean.*

8. Distribution of second moments. To study in more detail the distributions of m_2^* and m_2^0 the higher moments are computed and compared. Applying the formula for the distribution of rth moments we obtain, for m_2^0

$$\mu_1'(m_2^0) = \mu_2$$

$$\mu_2(m_2^0) = \frac{1}{n}(\mu_4 - \mu_2^2)$$

$$(14) \qquad \alpha_3(m_2^0) = \frac{\alpha_6 - 3\alpha_4 + 2}{\sqrt{n}\,(\alpha_4 - 1)^{3/2}}$$

$$\alpha_4(m_2^0) - 3 = \frac{1}{n}\left[\frac{\alpha_8 - 4\alpha_6 + 6\alpha_4 - 3}{(\alpha_4 - 1)^2} - 3\right]$$

etc.

For the distribution of m_2^*, we get

$$\mu_1'(m_2^*) = \mu_2 + \frac{\mu_3^2}{4\mu_2^2}$$

$$\mu_2(m_2^*) = \frac{1}{n}\left(\mu_4 - \mu_2^2 - \frac{\mu_3^2}{\mu_2}\right)$$

$$(15) \qquad \alpha_3(m_2^*) = \frac{\alpha_6 - 3\alpha_4 + 2 + 3\alpha_3^3 - 3\alpha_5\alpha_3 + 3\alpha_4\alpha_3^2 - \alpha_3^4}{\sqrt{n}\,(\alpha_4 - \alpha_3^2 - 1)^{3/2}}$$

$$\alpha_4(m_2^*) - 3 = \frac{1}{n}[(\alpha_8 - 4\alpha_6 + 6\alpha_4 - 3 + 12\alpha_5\alpha_3 - 6\alpha_3^2 - 4\alpha_7\alpha_3 + 6\alpha_6\alpha_3^2 - 12\alpha_4\alpha_3^2 + 4\alpha_3^4 - 4\alpha_3^3\alpha_5 + \alpha_4\alpha_3^4)(\alpha_4 - \alpha_3^2 - 1)^{-2} - 3]$$

etc.

Computing the ratios of α_3's, we have

$$(16) \qquad \frac{\alpha_3(m_2^*)}{\alpha_3(m_2^0)} = \left[1 - \frac{\alpha_3\{3(\alpha_5 - \alpha_3) - \alpha_3(3\alpha_4 - \alpha_3^2)\}}{\alpha_6 - 3\alpha_4 + 2}\right]\left(1 - \frac{\alpha_3^2}{\alpha_4 - 1}\right)^{-3/2}.$$

Similarly

$$(17) \qquad \frac{\alpha_4(m_2^*) - 3}{\alpha_4(m_2^0) - 3} = \left[1 - \frac{\alpha_3(4\alpha_7 + 6\alpha_4\alpha_3 + 4\alpha_3^2\alpha_5 + 12\alpha_3 - 12\alpha_5 - 6\alpha_6\alpha_3 - \alpha_3^3 - \alpha_3^3\alpha_4)}{\alpha_8 - 4\alpha_6 - 3\alpha_4^2 + 12\alpha_4 - 6}\right] \cdot \left(1 - \frac{\alpha_3^2}{\alpha_4 - 1}\right)^{-2}.$$

It is evident that when $\alpha_3 = 0$, the ratio in each case is unity. These ratios seem too involved to make any other general statements, but for particular types of populations these ratios in terms of the parameters are considerably simplified.

To illustrate this statement, consider

$$f_x = \frac{e^{-M}M^x}{x!}.$$

From the foregoing formulas we compute

$$\mu_1'(m_2^*) = M + \tfrac{1}{4}, \qquad \mu_1'(m_2^0) = M$$

$$\mu_2(m_2^*) = \frac{2M^2}{n}, \qquad \mu_2(m_2^0) = \frac{2M^2 + M}{n}$$

$$(18) \qquad \frac{\alpha_3(m_2^*)}{\alpha_3(m_2^0)} = \sqrt{\frac{2}{M}}\frac{(2M+1)^{5/2}}{8M^2 + 22M + 1}$$

$$(19) \qquad \frac{\alpha_4(m_2^*) - 3}{\alpha_4(m_2^0) - 3} = \frac{(12M^2 + 36M + 2)(2M + 1)^2}{M(48M^3 + 384M^2 + 112M + 1)}.$$

The minimum value of (18) is 0.71 for $M = 1.22$ and (18) is < 1 for $M > 0.31$. The minimum of (19) is 0.70 and is < 1 for $M > 0.62$. For the Poisson distribution, then, not only is the variance of m_2^* less than that of m_2^0, but at least as far as the first four moments are concerned, the distribution of m_2^* approaches normality more rapidly than does m_2^0 for all values of $M > 0.62$.

When one follows the same procedure for $\frac{1}{\Gamma(p)} x^{p-1} e^{-x}$ it is found that not only is the variance of m_2^* less than that of m_2^0, but as far as the first four moments are concerned, the distribution of m_2^* approaches normality more rapidly than does m_2^0, for values of $p > 0.7$.

In the case of higher moments, it seems desirable to solve the necessary equations in each particular case, since the equations are somewhat involved.

9. Examples. A few examples are exhibited to illustrate the foregoing ideas and to contrast with some of the other methods.

1. A sample of 1,000 is obtained with the following distribution

x:	0	1	2	3	4
f:	625	269	91	11	4

The hypothesis being tested is that the population is $f_x = \frac{e^{-M}M^x}{x!}$, with $M = 0.5$.

$\bar{x} = 0.5$ and therefore the mean does not differ from its expected value.

By using the m_2^* test, we compute $t = 2.06$. If m_2^* is distributed normally, the hypothesis is rejected at the 5% level. By using the m_2^0 test, we find $t = 1.45$, and therefore by this test the hypothesis is not rejected at the 5% level.

Applying the χ^2 test, we find that the hypothesis is not rejected at the 5% level.

2. We return now to the sample mentioned in the introduction.

Since the parameters in population A were found by fitting the first two moments, the tests will be made on the higher moments. From the definition of m_2^0 and m_2^* it is clear what is meant by m_3^0, m_3^*, m_4^0 and m_4^*.

Consider the discrepancy of third moments in standard units t as a function of h, the distance from the origin. It is easy to see that

$$t = (m_3' - \mu_3')/\sqrt{G},$$

where

$$G = \frac{1}{n}[\mu_6' - \mu_3'^2 - 6h(\mu_5' - \mu_3'\mu_2') + 3h^2(5\mu_4' - 3\mu_2'^2 - 2\mu_3'\mu_1') - 18h^3(\mu_3' - \mu_2'\mu_1') + 9h^4(\mu_2' - \mu_1'^2)].$$

For the m_3^0 test, $h = 139.288$. The value of h which minimizes the variance is a solution of $6(\mu_2' - \mu_1'^2)h^3 - 9(\mu_3' - \mu_2'\mu_1')h^2 + (5\mu_4' - 3\mu_2'^2 - 2\mu_3'\mu_1')h - (\mu_5' - \mu_3'\mu_2') = 0$, which, for this population is $h = 142.66$. Using these values and computing, we find, for the m_3^0 test, $t = 1.90$ and for the m_3^* test, $t = 1.95$.

Using the same methods applied to fourth moment tests, we obtain for the m_4^0 test, $h = 139.288$ and $t = 48.7$, and for the m_4^* test, $h = 143.73$ and $t = 52.4$.

The χ^2 test cannot be used here since the moments alone are given; furthermore there is some difficulty in interpreting it under these conditions.

In this particular example, the third moment test would not reject the hypothesis at the 1% level, while the fourth moment test would reject at that level.

3. Since population B is symmetric, it is known that the m_3^0 and m_3^* tests are identical; similarly for m_4^0 and m_4^*. For the m_3^* test, $t = 5.57$, which would reject the hypothesis at the 1% level. The fourth moment test would not be applied in practice.

The writer wishes to acknowledge his indebtedness to Professor P. S. Dwyer for counsel and guidance. He also wishes to thank Professors H. C. Carver and C. C. Craig for valuable suggestions.

REFERENCES

[1] S. S. Wilks, *Statistical Inference*. Princeton Univ. Press, 1937.

[2] *Tables for Statisticians*, Edwards Bros. Inc.

[3] N. P. Bertelsen, "On the compatability of frequency constants and on presumptive laws of error," *Skandenavisk Aktuar. tids.*, Vol. 10 (1927), pp. 129–156.

[4] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw Hill, 1937.

[5] K. Pearson, "Mathematical contributions to the theory of evolution, XIX, second supplement to a memoir on skew variation," *Roy. Soc. Phil. Trans.*, Series A, Vol. 216 (1916), pp. 429–457.

[6] C. C. Craig, "A new exposition and chart for the Pearson system of frequency curves," *Annals of Math. Stat.*, Vol. 7 (1936), pp. 16–28.

[7] G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge Univ. Press, 1934.