FUNDAMENTALS OF THE THEORY OF INVERSE SAMPLING[1]

By Ching-Lai Shen

Part I. Introduction[2]

Section I. Statistical Concepts of the Theory of Sampling

One of the chief objects in statistics is to form a judgment of a very large statistical universe, known as a parent population, by means of a study of a part or sample thereof, which is drawn at random. To make a complete survey of the parent population is sometimes impossible or impractical. For example, it is impossible to measure the heights of all adult persons in a country. It is impractical to test for infectious bacteria the whole body of water in a city reservoir. All that we can do is to obtain an unbiased sample. By an unbiased sample, we mean a sample in which each individual has an equal and independent chance to be included. From this chosen sample we attempt to draw some conclusion concerning the nature of the whole parent population in accordance with certain mathematical principles.

Now the sample which we choose is of course only one of the samples that can be possibly drawn from a given parent population. Suppose there is a population of s individuals from which we wish to choose a sample of r. It is clear that there exist ${}_sC_r$ such samples, each of which is equally likely to be chosen. Therefore these ${}_sC_r$ samples constitute the so-called distribution of samples. To describe from the statistical point of view the distribution of samples, we must find its mean, standard deviation, skewness, excess, and other higher characteristics. The first three are usually referred to as elementary statistical functions.

Suppose x_i be the variate (by which we mean the magnitude of a specified character of an individual to be measured) where $i = 1, 2, 3, \cdots s$; and z_j be the samples chosen from the parent population where $j = 1, 2, 3, \cdots {}_sC_r$. Then the ${}_sC_r$ samples, each consisting of r variables, will be formed after the following fashion:

$$z_1 = x_1 + x_2 + x_3 + \cdots + x_r$$

$$z_2 = x_2 + x_3 + x_4 + \cdots + x_{r+1}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$z_{\binom{s}{r}} = x_{s-r+1} + x_{s-r+2} + x_{s-r+3} + \cdots + x_s$$

---

[1] A dissertation submitted in partial fulfillment of the requirement for the degree of doctor of philosophy in the University of Michigan.

[2] The writer wishes to express his appreciation for the assistance Professor H. C. Carver has given him in making this study.

If we denote the nth moment of the parent population about its mean by

$$\bar{u}_{n:x} = \frac{\sum_{i=1}^{s} (x_i - M_x)^n}{s}$$

and the nth moment of the distribution of samples about its mean by

$$\bar{\mu}_{n:z} = \frac{\sum_{j=1}^{\binom{s}{r}} (z_j - M_z)^n}{\binom{s}{r}}$$

and if we then utilize the multinomial theorem, we may be able to express the sample moments in terms of the moments of the parent population:[3]

$$(1)\quad \begin{cases} M_z = rM_x \\ \bar{\mu}_{2:z} = 2!\left\{P_2 \dfrac{s\bar{\mu}_{2:x}}{2!}\right\} \\ \bar{\mu}_{3:z} = 3!\left\{P_3 \dfrac{s\bar{\mu}_{3:x}}{3!}\right\} \\ \bar{\mu}_{4:z} = 4!\left\{P_4 \dfrac{s\bar{\mu}_{4:x}}{4!} + \dfrac{P_2^2}{2!}\dfrac{s^2\bar{\mu}_{2:x}^2}{(2!)^2}\right\} \\ \bar{\mu}_{5:z} = 5!\left\{P_5 \dfrac{s\bar{\mu}_{5:x}}{5!} + P_3P_2 \dfrac{s^2\bar{\mu}_{3:x}\bar{\mu}_{2:x}}{3!\,2!}\right\} \\ \bar{\mu}_{6:z} = 6!\left\{P_6 \dfrac{s\bar{\mu}_{6:x}}{6!} + P_4P_2 \dfrac{s^2\bar{\mu}_{4:x}\bar{\mu}_{2:x}}{4!\,2!}\right. \end{cases}$$

$$\left. + \frac{P_3^2}{2!}\frac{s^2\bar{\mu}_{3:x}^2}{(3!)^2} + \frac{P_2^3}{3!}\frac{s^3\bar{\mu}_{2:x}^3}{(2!)^3}\right\}, \text{ etc.}$$

where P_n is obtained from the sampling polynomial $P_n(\rho)$ by writing ρ^i as ρ_i:

$$(2)\quad \begin{cases} P_1(\rho) = \rho \\ P_2(\rho) = \rho - \rho^2 \\ P_3(\rho) = \rho - 3\rho^2 + 2\rho^3 \\ P_4(\rho) = \rho - 7\rho^2 + 12\rho^3 - 6\rho^4 \\ P_5(\rho) = \rho - 15\rho^2 + 50\rho^3 - 60\rho^4 + 24\rho^5 \\ P_6(\rho) = \rho - 31\rho^2 + 180\rho^3 - 390\rho^4 + 360\rho^5 - 120\rho^6, \text{ etc.} \end{cases}$$

where

$$\rho_i = \frac{r(r-1)(r-2)\cdots(r-\overline{i-1})}{s(s-1)(s-2)\cdots(s-\overline{i-1})}$$

[3] Carver, H. C., *Annals of Mathematical Statistics*, Vol. I, No. I, pp. 106–107.

Section II. Frequency Curve of the Distribution of Samples

The frequency distribution of samples is usually less scattered than individual observations. In order to ascertain the manner of the distribution, we have access to the well-known Type A Curve of Charlier.[4]

$$F(t) = \phi(t) - \frac{c_3}{3!}\phi^{(3)}(t) + \frac{c_4}{4!}\phi^{(4)}(t) - \frac{c_5}{5!}\phi^{(5)}(t) + \cdots \tag{3}$$

where $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$

$$\begin{aligned} c_3 &= \alpha_3 \\ c_4 &= \alpha_4 - 3 \\ c_5 &= \alpha_5 - 10\alpha_3 \\ c_6 &= \alpha_6 - 15\alpha_4 + 30 \\ c_7 &= \alpha_7 - 21\alpha_5 + 10\alpha_3 \\ c_8 &= \alpha_8 - 28\alpha_6 + 210\alpha_4 - 315, \text{ etc.} \end{aligned}$$

This formula is a powerful tool for representing any frequency; but it is emphasized by more than one author[5] that the usefulness of such a series representation of a frequency distribution depends upon the rapidity of convergence, and the rapidity of convergence in turn depends upon the extent to which the function $\phi(t)$ is a fair approximation for $F(t)$. We shall not, however, discuss here the question of convergence. What we are interested in is to apply this series representation to the distribution of samples and see whether our numerical experimentation justifies the use of it.

TABLE I

Heights of 1000 Freshman Students

(Original Measurements Made to Nearest 0.1 in.)

Class	Frequency
58.5–60.4	2
60.5–62.4	13
62.5–64.4	76
64.5–66.4	167
66.5–68.4	335
68.5–70.4	264
70.5–72.4	106
72.5–74.4	29
74.5–76.4	7
76.5–78.4	1

[4] Camp, B. H., *The Mathematical Part of Elementary Statistics*, p. 226.

[5] Rietz, H. L., *Mathematical Statistics* p. 62.

Carver, H. C., Frequency Curves, *Handbook of Mathematical Statistics*, p. 115.

First of all, therefore, we take for our numerical example the heights of 1000 freshman students in the University of Michigan, as recorded in Table I, which are assumed to constitute our parent population.

From the above data we compute the first 6 moments as follows:

$$
\begin{aligned}
M_x &= 67.91 & & \\
\bar{\mu}_{2:x} &= 6.279{,}068 & \sigma_x &= 2.505{,}81 \\
\bar{\mu}_{3:x} &= 0.489{,}552 & \alpha_{3:x} &= 0.031{,}11 \\
\bar{\mu}_{4:x} &= 132.685{,}214 & \alpha_{4:x} &= 3.365{,}36 \\
\bar{\mu}_{5:x} &= 78.435{,}794 & \alpha_{5:x} &= 0.793{,}92 \\
\bar{\mu}_{6:x} &= 4574.080{,}554 & \alpha_{6:x} &= 18.476{,}43
\end{aligned}
$$

Now suppose from this parent population in which $s = 1000$, we wish to choose ${}_{1000}C_{100}$ samples, each consisting of 100 individuals. To characterize the distribution of these samples, we first make the following table:

TABLE II

Values of ρ_i and P_i for $s = 1000$, $r = 100$

$\rho_1 =$	.1
$\rho_2 =$	.009,909,909,91
$\rho_3 =$	.000,973,117,406
$\rho_4 =$	.000,094,676,417,6
$\rho_5 =$	.000,009,125,437,84
$\rho_6 =$	.000,000,871,272,959,5
$P_1 =$	.1
$P_2 =$	.090,090,090,09
$P_3 =$	.072,216,505,082
$P_4 =$	.041,739,980,994
$P_5 =$	−.005,454,352,918
$P_6 =$	−.065,789,272,230
$P_2^2 =$	.008,058,351,516
$P_2P_3 =$	.006,472,571,500
$P_2P_4 =$	.003,764,792,358
$P_2^3 =$	.000,715,593,194
$P_3^2 =$	.005,195,978,741

Substituting into formulae (1), we obtain the first six moments of the distribution of samples:

$$
\begin{aligned}
M_z &= 6791 & & \\
\bar{\mu}_{2:z} &= 565.621{,}622 & \sigma_z &= 23.782{,}8 \\
\bar{\mu}_{3:z} &= 35.353{,}734 & \alpha_{3:z} &= .002{,}628 \\
\bar{\mu}_{4:z} &= 958{,}720.852{,}854 & \alpha_{4:z} &= 2.996{,}679 \\
\bar{\mu}_{5:z} &= 198{,}538.702{,}142 & \alpha_{5:z} &= .026{,}093 \\
\bar{\mu}_{6:z} &= 2{,}704{,}514{,}780.791{,}465 & \alpha_{6:z} &= 14.945{,}539
\end{aligned}
$$

The coefficients of Charlier's Type A Curve turn out to be very small and rapidly decreasing:

$$\frac{c_3}{3!} = .000,438$$

$$\frac{c_4}{4!} = -.000,138$$

$$\frac{c_5}{5!} = -.000,016$$

$$\frac{c_6}{6!} = -.000,006$$

We therefore may be justified in considering this series representation of the sample distribution as converging rapidly to the normal curve. It may be interesting to note that even from a parent population which is very skew, the distribution of samples is nearly normal—as the following example will show:

TABLE III

Weights of 1000 Freshman Students

(Original Measurements Made to Nearest Pound)

Class	Frequency
85–	1
95–	8
105–	45
115–	132
125–	232
135–	244
145–	161
155–	97
165–	50
175–	16
185–	7
195–	3
205–	4

$M_x = 139.32$

$\bar{\mu}_{2:x} = 296.8343 \qquad \sigma_x = 17.228,87$

$\bar{\mu}_{3:x} = 3,230.802 \qquad \alpha_{3:x} = 0.631,74$

$\bar{\mu}_{4:x} = 351,180.14 \qquad \alpha_{4:x} = 3.985,67$

$\bar{\mu}_{5:x} = 11,811,480.5 \qquad \alpha_{5:x} = 7.780,71$

$\bar{\mu}_{6:x} = 886,585,271 \qquad \alpha_{6:x} = 33.898,36$

$$M_z = 13{,}932$$
$$\bar{\mu}_{2:z} = 26{,}741.828{,}829 \qquad \sigma_z = 163.529$$
$$\bar{\mu}_{3:z} = 233{,}317.229{,}045 \qquad \alpha_{3:z} = .05334$$
$$\bar{\mu}_{4:z} = 2{,}144{,}736{,}851.477{,}805 \qquad \alpha_{4:z} = 2.9991$$
$$\bar{\mu}_{5:z} = 62{,}008{,}368{,}279.121{,}883 \qquad \alpha_{5:z} = .53024$$
$$\bar{\mu}_{6:z} = 287{,}107{,}828{,}746{,}809.017 \qquad \alpha_{6:z} = 15.00633$$

$$\frac{c_3}{3!} = .008{,}89$$

$$\frac{c_4}{4!} = -.000{,}04$$

$$\frac{c_5}{5!} = -.000{,}03$$

$$\frac{c_6}{6!} = .000{,}03$$

Indeed the distribution of samples, in general, is very nearly normal irrespective of the law of distribution of the parent population. From the practical point of view, as Professor H. C. Carver has remarked, the parent population has little control over the shape of the distribution of the samples of r is fifty or greater and if S is at least ten times as large as r.[6]

Now as a numerical illustration of the theory of sampling I may, for example, choose at random 100 weights from the parent population of 1000 weights of freshman students, as recorded in Table III, with the aim of ascertaining the probability that the mean of this sample exceeds 142 pounds.

Since we define the mean of a sample simply as the average measurement of the r individuals in the sample, which in this case is 100, it therefore follows that the ordinary moments of the distribution of sample means differ from those of the distribution of samples in (1) only by a constant multiple of $1/r^k$ where k is the order of the moments concerned, while the standardized moments remain unchanged. Therefore in this problem, we have the mean of the sample means equal to 139.32 and the standard deviation equal to 1.63529. The average weight, 142 pounds, may be expressed in standard units as

$$t = \frac{z - M_z}{\sigma_z} = \frac{142 - 139.32}{1.63529} = 1.63885$$

In accordance with (3), the probability that the mean of the sample exceeds 142 pounds is therefore equal to

$$P = \int_{.63885}^{\infty} \left[\phi(t) - \frac{c_3}{3!}\phi^{(3)}(t) + \frac{c_4}{4!}\phi^{(4)}(t) - \frac{c_5}{5!}\phi^{(5)}(t) + \cdots \right] dt$$

[6] Carver, H. C., *Annals of Mathematical Statistics*, Vol. I, No. I, p. 112.

If we take the first term only, $P = \int_{1.63885}^{\infty} \phi(t)dt = .05062.$

If we take the first two terms, $P = \int_{1.63885}^{\infty} \phi(t)dt - .00889\phi^{(2)}(t)\Big]_{1.63885}^{\infty} =$ .05218.

If we take the first three terms,

$$P = \int_{1.63885}^{\infty} \phi(t)dt - .00889\phi^{(2)}(t)\Big]_{1.63885}^{\infty} + (-.000,04)\phi^{(3)}(t)\Big]_{1.63885}^{\infty} = .052182.$$

Section III. Pearsonian Types of Curves

Charlier's Type A Series is, however, not the only known analytic representation of a frequency distribution. There are Pearsonian Types of Curves, the characteristics of which I shall need to summarize briefly. These Pearsonian Types of Curves are essential to the later development of our theory.

The curves, suggested by certain geometrical properties of unimodal frequency distribution, are all obtained from the solution of the differential equation:

$$\frac{1}{y}\frac{dy}{dt} = \frac{a - t}{f(t)}$$

where $f(t)$ is assumed to be possibly expanded into a convergent power series, that is, $f(t) = b_0 + b_1 t + b_2 t^2 + \cdots$. When the first three terms of the power series are taken, the differential equation immediately takes the form of $\frac{1}{y}\frac{dy}{dt} = \frac{a - t}{b_0 + b_1 t + b_2 t^2}$. The parameters, a, b_0, b_1, b_2, may be expressed in terms of moments:[7]

$$a = -\frac{\alpha_3}{2(1 + 2\delta)} \qquad b_0 = \frac{2 + \delta}{2(1 + 2\delta)}$$

$$b_1 = \frac{\alpha_3}{2(1 + 2\delta)} \qquad b_2 = \frac{\delta}{2(1 + 2\delta)}$$

where

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

Based upon the difference in the nature of the roots of the equation $b_0 + b_1 t + b_2 t^2 = 0$, there have been derived thirteen types or curves. Of the particularly noteworthy ones, the normal curve and Type III may be mentioned. The criterion for the normal curve is $\alpha_3 = \delta = 0$; that for Type III is

[7] Carver, H. C., Frequency Curves, *Handbook of Mathematical Statistics*, p. 104.

$\delta = 0$ and $\alpha_3 \neq 0$. In order to fix the form in a particular case, we may refer to Pearson's Chart $\beta_1\beta_2$ Distribution[8] where

$$\beta_1 = \frac{\bar{\mu}_3^2}{\bar{\mu}_2^3} = \alpha_3^2, \qquad \beta_2 = \frac{\bar{\mu}_4}{\bar{\mu}_2^2} = \alpha_4,$$

and

$$K = \frac{b_1^2}{4b_0 b_2} = \frac{\beta_1(\beta_2 + 3)^2}{4(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)} = \frac{\alpha_3^2}{4\delta(2 + \delta)},$$

or to Elderton's *Frequency Curves and Correlation*.[9]

Section IV. The Inverse Sampling, Our Problem

It is now our problem to study the theory of inverse sampling, by which we mean that given the characteristics of a single sample drawn at random from a parent population, we wish to ascertain the probability that the corresponding characteristics of that parent population do not differ from those observed in the sample by more than a specified amount. To illustrate, suppose we are interested in knowing the average height of 1000 freshman students to which reference has already been made. Due to the fact that it takes too much time or is otherwise impractical to measure all of them so as to obtain the true average, we select at random one hundred of them and measure the heights of these one hundred individuals. Suppose the mean, the standard deviation, and the skewness of this sample of one hundred are computed and they are as follows:

$$M = 67.99$$

$$\sigma = 2.327$$

$$\alpha_3 = -.12299$$

Now assuming that the true mean of the entire 1000 heights is unknown, let us find the probability that the true mean of this parent population lies between $M_x = a$ and $M_x = b$ by what we know of the characteristics of the observed sample of one hundred as recorded above. It is clear that if we can obtain an equation, $y = f(M_x)$, of the frequency curve associated with the distribution of hypothetical means of this parent population, we shall be able to ascertain the probability we desire by evaluating the following integral expression:

$$P = \frac{\int_a^b f(M_x)dM_x}{\int_{-\infty}^{\infty} f(M_x)dM_x}$$

[8] Pearson, K., *Tables for Statisticians and Biometricians*, Vol. II, front page.

[9] Elderton, W. P., *Frequency Curves and Correlation*, Table VI, opposite p. 46.

In the same way we can find the probability that the standard deviation of the parent population lies between two definite limits or that the skewness of the parent population lies between two definite limits.

Our procedure will therefore be as follows: First, assuming the *a priori* existence of a continuous sequence of hypothetical means of the parent population, we investigate the relation between the distribution of these hypothetical means of the parent population and the distribution of sample means. If such a relation exists, we shall be able to find an expression for the most probable value of the parent mean. Assuming the most probable value of the parent mean to be the true mean of the parent population, we shall obtain an expression for the most probable value of the standard deviation of the parent population. Then it will be possible for us to express the frequency curve associated with the distribution of hypothetical means of the parent population in the form of $f(M_x)$. Similarly we may find the frequency functions associated with the standard deviation and skewness of the parent population.

Before leaving this section, it is perhaps not out of place to say a word about the connection of this theory of inverse sampling with Bayes's Theorem. The theory of inverse sampling (which deals essentially with the problem of judging the nature of a whole by observation of a part of it) belongs to the domain of inductive probability, or inverse probability, upon which Bayes's Theorem was founded. In order to solve a problem of inductive probability, it is necessary to postulate the *a priori* existence of the causes from which an event takes place, which, in our case, is the hypothetical means of the parent population.

This *a priori* hypothesis which gives rise to Bayes's Theorem has been viewed with suspicion by a number of mathematical statisticians. For example, the theorem has been called into question by such mathematicians as Bing, Venn, Chrystal, and others, including several now living. But so far as the present writer is aware, no definite conclusion has been reached. It is true that on the one hand Bayes's Theorem has not been rigidly demonstrated and proved by logic; but on the other hand the process of generalization from observational data is justified within the limits of ordinary practical application. One who holds Bayes's Theorem strongly may even say that the *a priori* hypothesis is absolutely necessary to scientific inferences. Concerning this controversy, Pearson takes a liberal point of view: "I hold this theorem [Bayes's Theorem] not as rigidly demonstrated, but I think with Edgeworth that the hypothesis of the equal distribution of ignorance is within the limits of practical life justified by experience of statistical ratios, which *a priori* are unknown · · · ."[10] He has further remarked that "the practical man · · · will accept the results of inverse probability of Bayes-Laplace brand till better are forthcoming."[11] Using

---

[10] Pearson, K., On the Influence of Past Experience on Future Expectation, *Philosophical Magazine*, Vol. 13, Jan.–June, 1907, p. 366.

[11] Pearson, K., The Fundamental Problem of Practical Statistics, *Biometrika*, Vol. 13, 1920–21, p. 3.

Pearson's viewpoint, we shall proceed with our problem by postulating *a priori* the existence of hypothetical means of the parent population from which our sample is drawn.

Part II. Fundamental Relation between the Moments of the Distribution of Sampling Means and the Moments of the Distribution of the Hypothetical Means Associated with the Parent Population

The characteristics of the distribution of sample means, as we have pointed out in Part I, Section II, differ from those of the sample distribution only by a constant multiple of $(1/r)^k$ where k is the order of the moments concerned. We may write down the first six moments of the distribution of sample means:

$$
(4)\quad \begin{cases}
M_{z_x} = M_x \\
\bar{\mu}_{2:z_x} = 2!\,\dfrac{s}{r^2}\left\{P_2\,\dfrac{\bar{\mu}_{2:x}}{2!}\right\} \\
\bar{\mu}_{3:z_x} = 3!\,\dfrac{s}{r^3}\left\{P_3\,\dfrac{\bar{\mu}_{3:x}}{3!}\right\} \\
\bar{\mu}_{4:z_x} = 4!\,\dfrac{s}{r^4}\left\{P_4\,\dfrac{\bar{\mu}_{4:x}}{4!} + \dfrac{P_2^2}{2!}\,\dfrac{s\,\bar{\mu}_{2:x}^2}{(2!)^2}\right\} \\
\bar{\mu}_{5:z_x} = 5!\,\dfrac{s}{r^5}\left\{P_5\,\dfrac{\bar{\mu}_{5:x}}{5!} + P_3P_2\,\dfrac{s\,\bar{\mu}_{3:x}\,\bar{\mu}_{2:x}}{3!\,2!}\right\} \\
\bar{\mu}_{6:z_x} = 6!\,\dfrac{s}{r^6}\left\{P_6\,\dfrac{\bar{u}_{6:x}}{6!} + P_4P_2\,\dfrac{s\,\bar{\mu}_{4:x}\,\bar{\mu}_{2:x}}{4!\,2!} + \dfrac{P_3^2}{2!}\,\dfrac{s\,\bar{\mu}_{3:x}^2}{(3!)^2} + \dfrac{P_2^3}{3!}\,\dfrac{s^2\,\bar{\mu}_{2:x}^3}{(2!)^3}\right\}
\end{cases}
$$

From these we immediately obtain

$$
(5)\quad \begin{cases}
M_{z_x} = M_x \\
\sigma_{z_x} = \sqrt{\dfrac{s-r}{r(s-1)}}\,\sigma_x \\
\alpha_{3:z_x} = \dfrac{s-2r}{s-2}\sqrt{\dfrac{s-1}{r(s-r)}}\,\alpha_{3:x} \\
\alpha_{4:z_x} - 3 = \dfrac{(s-1)(s^2+s-6rs+6r^2)}{r(s-r)(s-2)(s-3)}\{\alpha_{4:x}-3\} \\
\qquad - \dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}, \text{ etc.}
\end{cases}
$$

If our parent population is infinite, which is a special case by allowing $s \to \infty$, then we have

$$(6)\qquad \begin{cases} M_{z_x} = M_x \\ \sigma_{z_x} = \dfrac{1}{\sqrt{r}}\,\sigma_{3:x} \\ \alpha_{3:z_x} = \dfrac{1}{\sqrt{r}}\,\alpha_{3:x} \\ \alpha_{4:z_x} - 3 = \dfrac{1}{r}\,(\alpha_{4:x} - 3), \text{ etc.} \end{cases}$$

Let us now define $f(t)$ as a frequency function of the distribution of sample means z_x in standard units, i.e.,

$$(7)\qquad t = \frac{Z_x - M_{z_x}}{\sigma_{z_x}}$$

Denoting the observed mean of a given sample by m_1 and making proper substitutions of (5), we obtain

$$(8)\qquad t = \frac{m_1 - M_x}{\sigma_{z_x}} = \frac{m_1 - M_x}{\sqrt{\dfrac{s-r}{r(s-1)}}\,\sigma_x}$$

It is clear that if we hold s, r and σ_x constant and let M_x vary, then t is a function of M_x only and consequently $f(t)$ becomes a function of M_x.

Suppose now $M_x^{(1)}, M_x^{(2)}, M_x^{(3)}, \cdots$ be a continuous sequence of hypothetical means, which M_x has an equal chance to assume. These hypothetical means will certainly lie in a linear interval between their natural limits. Then the probability that M_x lies between $M_x \pm \frac{1}{2}\,dM_x$ is $f(t)dM_x$. Therefore, to obtain the probability that M_x lies in the interval $M_x^{(i)} \leq M_x \leq M_x^{(i+1)}$, it is only necessary to carry out the integration of this expression:

$$(9)\qquad \int_{M_x^{(i)}}^{M_x^{(i+1)}} f(t)\,dM_x$$

There is no question as regards the existence of this integral in case of an infinite parent population. As for a finite population, we may still use this continuous function as an interpolation function to the true discontinuous function.

Let us now define $P(t)$ as the probability function for which the hypothetical mean of the parent population falls within certain specified limits. Considering

$\mu_{n:p}$ as the nth moment of this probability function about a fixed point, we will have the following relation:

$$\mu_{n:p} = \frac{\int_{-l}^{l} M_x^n f(t) dM_x}{\int_{-l}^{l} f(t) dM_x} \tag{10}$$

where l and $-l$ are their natural limits.

Since from (8), $M_x = m_1 - \sigma_{z_x} t$, then after substitution, we obtain

$$\begin{aligned} \mu_{n:p} &= \frac{\sigma_{z_x} \int_{\frac{m_1 - l}{\sigma_{z_x}}}^{\frac{m_1 + l}{\sigma_{z_x}}} (m_1 - \sigma_{z_x} t)^n f(t) dt}{\sigma_{z_x} \int_{\frac{m_1 - l}{\sigma_{z_x}}}^{\frac{m_1 + l}{\sigma_{z_x}}} f(t) dt} \\ &= \int_{\frac{m_1 - l}{\sigma_{z_x}}}^{\frac{m_1 + l}{\sigma_{z_x}}} (m_1 - \sigma_{z_x} t)^n f(t) dt \\ &= m_1^n - \binom{n}{1} m_1^{n-1} \bar{\mu}_{1:z_x} + \binom{n}{2} m_1^{n-2} \bar{\mu}_{2:z_x} - \binom{n}{3} m_1^{n-3} \bar{\mu}_{3:z_x} \\ &\quad + \cdots + (-1)^n \binom{n}{n} \bar{\mu}_{n:z_x} \end{aligned} \tag{11}$$

$$\begin{cases} \mu_{1:p} = M_p = m_1 \\ \mu_{2:p} = m_1^2 + \bar{\mu}_{2:z_x} \\ \mu_{3p} = m_1^3 + 3m_1 \bar{\mu}_{2:z_x} - \bar{\mu}_{3:z_x} \\ \mu_{4:p} = m_1^4 + 6m_1^2 \bar{\mu}_{2:z_x} - 4m_1 \bar{\mu}_{3:z_x} + \bar{\mu}_{4:z_x}, \text{ etc.} \end{cases} \tag{12}$$

The first relation $M_p = m_1$ is important because it shows that the mean of the hypothetical means of the parent population is equal to the mean of the observed sample drawn from it. To state this in a theorem, we will have

Theorem I. The expected value of a parent mean is equal to the mean of an observed sample chosen from the parent population.

We now wish to express the moments of the probability function about its mean in terms of the moments of sample distribution. In general, the nth moment of any frequency distribution about its mean, $\bar{\mu}_n$, can be expressed in terms of its moments about a fixed point after the following fashion:

$$\bar{\mu}_n = \mu_n - \binom{n}{1} M \mu_{n-1} + \binom{n}{2} M^2 \mu_{n-2} - \cdots + (-1)^n \binom{n}{n} M^n. \tag{13}$$

Therefore when we substitute (11) into (13) we obtain

$$
\begin{aligned}
\bar{\mu}_{n:p} = m_1^n &- \binom{n}{1} m_1^{n-1}\, \bar{\mu}_{1:z_x} + \binom{n}{2} m_1^{n-2}\, \bar{\mu}_{2:z_x} - \cdots \\
&+ (-1)^{n-3}\binom{n}{n-3} m_1^3\, \bar{\mu}_{n-3:z_x} + (-1)^{n-2}\binom{n}{n-2} m_1^2\, \bar{\mu}_{n-2:z_x} \\
&+ (-1)^{n-1}\binom{n}{n-1} m_1\, \bar{\mu}_{n-1:z_x} + (-1)^n \binom{n}{n} \bar{\mu}_{n:z_x} \\
- m_1\binom{n}{1}\Bigg[m_1^{n-1} &- \binom{n-1}{1} m_1^{n-2}\, \bar{\mu}_{1:z_x} + \binom{n-1}{2} m_1^{n-3}\, \bar{\mu}_{2:z_x} - \cdots \\
&+ (-1)^{n-3}\binom{n-1}{n-3} m_1^2\, \bar{\mu}_{n-3:z_x} + (-1)^{n-2}\binom{n-1}{n-2} m_1\, \bar{\mu}_{n-2:z_x} \\
&+ (-1)^{n-1}\binom{n-1}{n-1} \bar{\mu}_{n-1:z_x}\Bigg] \\
+ m_1^2\binom{n}{2}\Bigg[m_1^{n-2} &- \binom{n-2}{1} m_1^{n-3}\, \bar{\mu}_{1:z_x} + \binom{n-2}{2} m_1^{n-4}\, \bar{\mu}_{2:z_x} - \cdots \\
&+ (-1)^{n-3}\binom{n-2}{n-3} m_1\, \bar{\mu}_{n-3:z_x} + (-1)^{n-2}\binom{n-2}{n-2} \bar{\mu}_{n-2:z_x}\Bigg] \\
- m_1^3\binom{n}{3}\Bigg[m_1^{n-3} &- \binom{n-3}{1} m_1^{n-4}\, \bar{\mu}_{1:z_x} + \binom{n-3}{2} m_1^{n-5}\, \bar{\mu}_{2:z_x} - \cdots \\
&+ (-1)^{n-3}\binom{n-3}{n-3} \bar{\mu}_{n-3:z_x}\Bigg] \\
+ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&+ (-1)^{n-1} m_1^{n-1}\binom{n}{n-1}\left[m_1 - \binom{1}{1}\bar{\mu}_{1:z_x}\right] \\
&+ (-1)^n m_1^n \binom{n}{n}
\end{aligned}
$$

Adding vertically each column, we obtain

$$
\begin{aligned}
\bar{\mu}_{x:p} = m_1^n &\left[\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \binom{n}{4} - \cdots + (-1)^n\binom{n}{n}\right] \\
- m_1^{n-1}\, \bar{\mu}_{1:z_x} &\Bigg[\binom{n}{1}\binom{n}{0} - \binom{n-1}{1}\binom{n}{1} + \binom{n-2}{1}\binom{n}{2} - \binom{n-3}{1}\binom{n}{3} \\
&+ \binom{n-4}{1}\binom{n}{4} - \cdots + (-1)^{n-1}\binom{n}{n-1}\Bigg]
\end{aligned}
$$

$$+ m_1^{n-2}\,\bar{\mu}_{2:zx}\left[\binom{n}{2}\binom{n}{0} - \binom{n-1}{2}\binom{n}{1} + \binom{n-2}{2}\binom{n}{2} - \binom{n-3}{2}\binom{n}{3} + \binom{n-4}{2}\binom{n}{4} - \cdots + (-1)^{n-2}\binom{n}{n-2}\right]$$

$$-\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ (-1)^{n-1} m_1\,\bar{\mu}_{n-1:zx}\left[\binom{n}{n-1}\binom{n}{0} - \binom{n-1}{n-1}\binom{n}{1}\right]$$

$$+ (-1)^n\,\bar{\mu}_{n:zx}$$

The first row of the above expression is equal to $m_1^n(1 - 1)^n = 0$; the second row is equal to

$$- m_1^{n-1}\,\bar{\mu}_{1:zx}\left[\frac{n!}{0!\,n!}\cdot\frac{n!}{1!\,(n-1)!} - \frac{n!}{1!\,(n-1)!}\cdot\frac{(n-1)!}{1!\,(n-2)!} + \frac{n!}{2!\,(n-2)!}\,\frac{(n-2)!}{1!\,(n-3)!} - \cdots + (-1)^n\frac{n!}{(n-1)!\,1!}\right]$$

$$= -m_1^{n-1}\bar{\mu}_{1:zx}\,\frac{n!}{1!}\left[\frac{1}{0!\,(n-1)!} - \frac{1}{1!\,(n-2)!} + \frac{1}{2!\,(n-3)!} - \cdots + (-1)^n\frac{1}{(n-1)!\,0!}\right]$$

$$= -m_1^{n-1}\bar{\mu}_{1:zx}\,\frac{n}{1!}\,[1 - {}_{n-1}C_1 + {}_{n-1}C_2 - \cdots + (-1)^n\,{}_{n-1}C_{n-1}]$$

$$= -m_1^{n-1}\bar{\mu}_{1:zx}\,\frac{n}{1!}\,(1-1)^{n-1} = 0;$$

the third row is equal to

$$m_1^{n-2}\bar{\mu}_{2:zx}\left[\frac{n!}{0!\,n!}\cdot\frac{n!}{2!\,(n-2)!} - \frac{n!}{1!\,(n-1)!}\cdot\frac{(n-1)!}{2!\,(n-3)!} + \frac{n!}{2!\,(n-2)!}\cdot\frac{(n-2)!}{2!\,(n-4)!} - \cdots + (-1)^{n-1}\frac{n!}{(n-2)!\,2!}\right]$$

$$= m_1^{n-2}\bar{\mu}_{2:zx}\,\frac{n(n-1)}{2!}\,[1 - {}_{n-2}C_1 + {}_{n-2}C_2 - \cdots + (-1)^{n-1}\,{}_{n-2}C_{n-2}]$$

$$= m_1^{n-2}\bar{\mu}_{2:zx}\,\frac{n(n-1)}{2!}\,(1-1)^{n-2} = 0;$$

and similarly all the other rows turn out to be zero except the last one which is equal to $(-1)^n\bar{\mu}_{n:zx}$

$$\bar{\mu}_{n:p} = (-1)^n\bar{\mu}_{n:zx} \tag{14}$$

This may be rewritten as

$$(15) \qquad \begin{cases} \bar{\mu}_{2n:p} = \bar{\mu}_{2n:z_x} \\ \bar{\mu}_{2n+1:p} = -\bar{\mu}_{2n+1:z_x} \end{cases}$$

or in standard units

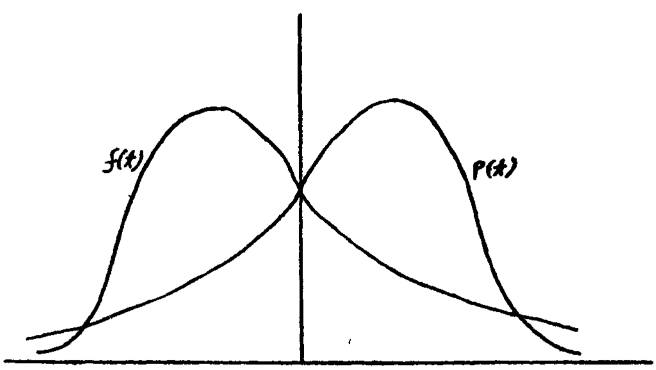
$$\begin{cases} \alpha_{2n:p} = \alpha_{2n:z_x} \\ \alpha_{2n+1:p} = -\alpha_{2n+1:z_x} \end{cases}$$

The results[12] of (15) are important and fundamental because they establish the relation between the Theory of Inverse Sampling and the Theory of Sampling. Therefore we may formulate the following theorems:

Theorem II. The even moments of the distribution of the hypothetical means of a parent population about its mean are equal to the corresponding even moments of the distribution of the sample means about the mean.

Theorem III. The odd moments of the distribution of the hypothetical means of a parent population about its mean are equal to the negative of the corresponding odd moments of the distribution of the sample means about the mean.

Since the even moments of the two distributions are the same, while the odd moments differ only in sign, it is evident that for symmetrical distributions, the two curves $f(t)$ and $P(t)$ are exactly identical, because in a symmetrical distribution all the odd moments about the mean are bound to vanish. In case of nonsymmetrical distributions, the curve $P(t)$ is nothing but a vertical reflection of the curve $f(t)$ as shown in the figure:



In other words, if $f(t)$, for instance, assumes Pearson's Type III Function, then $P(t)$ also assumes Pearson's Type III Function except that their skewness is different in sign though equal numerically. We therefore state our theorem as follows:

[12] So far as the writer is aware, these theorems were first developed by Professor H. C. Carver.

Theorem IV. The curves for the distribution of the hypothetical means of the parent population and the curve for the distribution of the means of the sample obtained from the parent population are symmetrically situated and one is a vertical reflection of the other.

Part III. Inverse Sampling Associated with a Normal Parent Population

We shall be concerned in this part of our discussion with a normal parent population. In accordance with the characteristics of a normal parent population we wish to investigate the most probable values of its mean and variance, thereby obtaining the distributions of the hypothetical means and variances of the parent population.

Section I. Most Probable Value of the Mean of the Parent Population

In Part I, Section III, we have mentioned Pearsonian Types of Frequency Curves whose differential equation is

$$\frac{1}{t}\frac{dy}{dt} = \frac{a - t}{b_0 + b_1 t + b_2 t^2}.$$

It is clear that the mode of these curves is at $t = a$, provided the mode exists. But to recapitulate:

$$a = \frac{-\alpha_3}{2(1 + 2\delta)},$$

where

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

consequently for the mode of the distribution of sample means, we have

$$t = \frac{-\alpha_{3:z_x}}{2(1 + 2\delta_{z_x})}, \tag{16}$$

where

$$\delta_{z_x} = \frac{2\alpha_{4:z_x} - 3\alpha_{3:z_x}^2 - 6}{\alpha_{4:z_x} + 3} \tag{17}$$

$$= \frac{2(s-1)(s-2)(s^2+s-6rs+6r^2)(\alpha_{4:x}-3) - 12s(r-1)(s-2)(s-r-1) - 3(s-1)(s-3)(s-2r)^2\alpha_{3:x}^2}{(s-2)\{(s-1)(s^2+s-6rs+6r^2)(\alpha_{4:x}-3) - 6s(r-1)(s-r-1) + 6r(s-r)(s-2)(s-3)\}}$$

$$\therefore\ t = \frac{z_x - M_{z_x}}{\sigma_{z_x}} = \frac{z_x - M_x}{\sqrt{\dfrac{s-r}{r(s-1)}}\,\sigma_x} = \frac{-\dfrac{s-2r}{s-2}\sqrt{\dfrac{s-1}{r(s-r)}}\,\alpha_{3:x}}{2(1 + 2\delta_{z_x})} \tag{18}$$

Now according to Theorem IV, the mode of the probability function $P(t)$ is situated symmetrically with respect to the mode of the frequency function $f(t)$ of the distribution of sample means; hence, for the mode of the probability function of hypothetical means of the parent population, we have

$$(19) \qquad t = \frac{m_1 - M_x}{\sqrt{\frac{s-r}{r(s-1)}}\,\sigma_x} = \frac{\frac{s-2r}{s-2}\sqrt{\frac{s-1}{r(s-r)}}\,\alpha_{3:x}}{2(1+2\delta_{z_x})}$$

where δ_{z_x} remains unchanged because it is a function of $\alpha^2_{3:z_x}$ and $\alpha_{4:z_x}$, each being always positive.

Solving for M_x, which will now be the most probable value of the mean of the parent population and hence denoted by $\hat{M}_x$, we have

$$(20) \qquad \hat{M}_x = m_1 - \frac{s-2r}{r(s-2)}\,\frac{\sigma_x \alpha_{3:x}}{2(1+2\delta_{z_x})}$$

It is interesting to note that if $s = 2r$, this expression yields $\hat{M}_x = m_1$, irrespective of the law of distribution of the parent population provided only that δ_{z_x} is not exactly equal to $-\frac{1}{2}$. But since the Pearson's function is used for graduation, one should not fail to see that the mode so obtained gives only an approximation to the true mode. Therefore we state a theorem as follows:

Theorem V. If a sample is composed of one-half of the variates of the parent population from which the sample is chosen, then the best approximated 'most probable value' of the mean of the parent population is equal to the mean of the observed sample provided only that δ_{z_x} is not exactly equal to $-\frac{1}{2}$.

It is further observed that if $\alpha_{3:x} = 0$ but $\delta_{z_x} \neq -\frac{1}{2}$, then the expression (20) will likewise yield $\hat{M}_x = m_1$. But $\alpha_{3:x} = 0$ implies that the frequency curve of the parent population is symmetrical. Hence

Theorem VI. For any symmetrical curves associated with the distribution of the parent population, the best approximated 'most probable value' of the mean of the parent population is equal to the mean of the observed sample provided δ_{z_x} is not exactly equal to $-\frac{1}{2}$.

But we will investigate further the most probable value of the mean of a normal parent population, and we know that in a normal distribution the moments bear the following relation:[13]

$$(21) \qquad \begin{cases} \alpha_{2n} = \dfrac{(2n)!}{2^n\, n!} \\ \alpha_{2n+1} = 0 \end{cases}$$

[13] Carver, H. C., Frequency Curves, *Handbook of Mathematical Statistics*, p. 97.

i.e., $\alpha_3 = 0$
$\alpha_4 = 3$
$\alpha_5 = 0$
$\alpha_6 = 15$
$\alpha_7 = 0$
$\alpha_8 = 105$
etc.

Consequently for a normal parent population the α_{z_x} function in (17) is immediately reduced to

$$\delta_{z_x} = \frac{2s(r-1)(s-r-1)}{s(r-1)(s-r-1) - r(s-r)(s-2)(s-3)} \tag{22}$$

Let us, first of all, investigate the possibility that this expression will be exactly equal to $-\frac{1}{2}$ for positive integral values of r and s.

Suppose we set

$$\frac{2s(r-1)(s-r-1)}{s(r-1)(s-r-1) - r(s-r)(s-2)(s-3)} = -\frac{1}{2}$$

and solve r in terms of s. Thus we obtain

$$r = \frac{s}{2} \pm \frac{\sqrt{s^2(s^2-10s+6)^2 + 20s(s-1)(s^2-10s+6)}}{2(s^2-10s+6)} \tag{23}$$

If $s \geqq 10$, then the second term on the right side is positive. As it is absurd that r should be greater than s, therefore the positive sign of the double sign should not be taken. Then, as the second term is obviously greater than $\frac{s}{2}$, the right member will be negative. Since r cannot be negative, no positive integral values of r and s, for which $s \geqq r$, can satisfy (23). For $s < 10$, there are only nine positive integers; and direct substitution of each will tell us that only when $s = 1, 2$, or 3, r is a positive integer which is either 1 or 2. As these are trifle cases because a parent population can never be so small, we may safely say that for a normal parent population

$$\hat{M}_x = m_1 \tag{24}$$

Theorem VII. For a normal parent population, the best approximated 'most probable value' of the mean of the parent population is equal to the mean of the observed sample from it.

For an infinite parent population, i.e., $s \to \infty$ (20) yields on reduction

$$\hat{M}_x = m_1 - \frac{1}{r}\frac{\sigma_x \alpha_{3:x}}{2(1+2\delta_{z_x})} \tag{25}$$

where

$$\delta_{z_x} = \frac{2(\alpha_{4:x}-3) - 3\alpha_{3:x}^2}{(\alpha_{4:x}-3) - 6r} \text{ [(from 17)}$$

Formula (25) yields immediately $\hat{M}_x = m_1$ if $\alpha_{3:x} = 0$ and $\delta_{z_x} \neq -\frac{1}{2}$. For a normal parent population $\delta_{z_x} = 0$. Hence Theorem VI and Theorem VII both hold for the infinite case.

Section II. Most Probable Value of the Standard Deviation of the Parent Population

To find the most probable value of the standard deviation of the parent population, we shall assume the mean of the parent population to be the best approximated 'most probable value' of the mean, which we have obtained in the preceding section. This assumption is necessary since we do not know the true mean of the parent population.

Now, to start with, we shall consider ${}_sC_r$ possible samples, each consisting of r variables. The second moment of each sample computed about the best approximated 'most probable value' of the mean of the parent population may be written as

$$z_1 = \frac{1}{r}\{(x_1 - m_1)^2 + (x_2 - m_1)^2 + (x_3 - m_1)^2 + \cdots + (x_r - m_1)^2\}$$

$$z_2 = \frac{1}{r}\{(x_2 - m_1)^2 + (x_3 - m_1)^2 + (x_4 - m_1)^2 + \cdots + (x_{r+1} - m_1)^2\}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$z_{\binom{s}{r}} = \frac{1}{r}\{(x_{s-r+1} - m_1)^2 + (x_{s-r+2} - m_1)^2 + (x_{s-r+3} - m_1)^2 + \cdots + (x_s - m_1)^2\}$$

If we write $(x_i - m_1)^2 = y_i$, it is clear that the above may be considered as a distribution of sample means drawn from a parent population $y_1, y_2, y_3 \cdots y_s$; and consequently

$$(26)\quad \begin{cases} M_{z_y} = M_y \\ \sigma_{z_y} = \sigma_y \sqrt{\dfrac{s-r}{r(s-1)}} \\ \alpha_{3:z_y} = \dfrac{s-2r}{s-2}\sqrt{\dfrac{s-1}{r(s-r)}}\,\alpha_{3:y} \\ \alpha_{4:z_y} - 3 = \dfrac{(s-1)(s^2+s-6rs+6r^2)}{r(s-r)\,s-2)(s-3)}\{\alpha_{4:y} - 3\} \\ \qquad - \dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}. \end{cases}$$

Now the nth moment of y about a fixed point may be written as

$$
\begin{aligned}
\mu_{n:y} &= \frac{1}{N}\sum y^n = \frac{1}{N}\sum (x - m_1)^{2n} \\
&= \frac{1}{N}\sum \{(x - M_x) + (M_x - m_1)\}^{2n} \\
&= \bar{\mu}_{2n:x} + \binom{2n}{1}\bar{\mu}_{2n-1:x}(M_x - m_1) \\
&+ \binom{2n}{2}\bar{\mu}_{2n-2:x}(M_x - m_1)^2 + \binom{2n}{3}\bar{\mu}_{2n-3:x}(M_x - m_1)^3 \\
&+ \binom{2n}{4}\bar{\mu}_{2n-4:x}(M_x - m_1)^4 + \cdots + (M_x - m_1)^{2n}.
\end{aligned} \tag{27}
$$

On the assumption that our parent population is normally distributed and due to the fact that in a normally distributed function

$$\alpha_{2n} = \frac{(2n)!}{2^n\, n!}, \quad \text{and} \quad \alpha_{2n+1} = 0 \qquad \text{[See (21)],}$$

the expression (27) immediately takes this form:

$$
\begin{aligned}
\mu_{n:y} &= \frac{2n!}{2^n \cdot n!}\sigma_x^{2n} + \binom{2n}{2}\frac{(2n-2)!}{2^{n-1}(n-1)!}(M_x - m_1)^2\sigma_x^{2n-2} \\
&+ \binom{2n}{4}\frac{(2n-4)!}{2^{n-2}(n-2)!}(M_x - m_1)^4\sigma_x^{2n-4} + \cdots + (M_x - m_1)^{2n}.
\end{aligned} \tag{28}
$$

Imposing the condition mentioned at the beginning of this section (i.e., M_x assumes its best approximated 'most probable value' m_1), then all the terms drop out except the first one. Hence, as a final form, we have

$$\mu_{n:y} = \frac{2n!}{2^n \cdot n!}\sigma_x^{2n} \tag{29}$$

$$
\begin{cases}
\mu_{1:y} = M_y = \sigma_x^2 \\
\mu_{2:y} = 3\sigma_x^4 \\
\mu_{3:y} = 15\sigma_x^6 \\
\mu_{4:y} = 105\sigma_x^8 \\
\quad \text{etc.}
\end{cases} \tag{30}
$$

It follows that the kth moment of y about its mean will be

$$\bar{\mu}_{k:y} = \frac{\sum (y - M_y)^k}{N} = \mu_{k:y} - \binom{k}{1} \mu_{k-1:y} M_y + \binom{k}{2} \mu_{k-2:y} M_y^2 - \cdots + (-1)^k \binom{k}{k} M_y^k$$

$$= \frac{2k!}{2^k \cdot k!} \sigma_x^{2k} - \binom{k}{1} \frac{(2k-2)!}{2^{k-1}(k-1)!} \sigma_x^{2k} + \binom{k}{2} \frac{(2k-4)!}{2^{k-2}(k-2)!} \sigma_x^{2k} - \cdots + (-1)^k \sigma_x^{2k}$$

(31) $$= \sigma_x^{2k} \left[\frac{2k!}{2^k \cdot k!} - \binom{k}{1} \frac{(2k-2)!}{2^{k-1}(k-1)!} + \binom{k}{2} \frac{(2k-4)}{2^{k-2}(k-2)!} - \cdots + (-1)^k \right]$$

$$= \sigma_x^{2k} \frac{2k!}{2^k \cdot k!} \left[1 - \frac{k}{1!(2k-1)} + \frac{k(k-1)}{2!(2k-1)(2k-3)} - \frac{k(k-1)(k-2)}{3!(2k-1)(2k-3)(2k-5)} + \cdots + (-1)^k \frac{k!}{k!(2k-1)(2k-3)(2k-5)\cdots(3)\cdot(1)} \right]$$

(32) $$\begin{cases} \bar{\mu}_{1:y} = 0 \\ \bar{\mu}_{2:y} = 2\sigma_x^4 \\ \bar{\mu}_{3:y} = 8\sigma_x^6 \\ \bar{\mu}_{4:y} = 60\sigma_x^8 \\ \bar{\mu}_{5:y} = 544\sigma_x^{10} \\ \bar{\mu}_{6:y} = 6040\,\sigma_x^{12} \\ \quad \text{etc.} \end{cases}$$

And therefore we obtain

(33) $$\begin{cases} \alpha_{3:y} = 2\sqrt{2} \\ \alpha_{4:y} = 15 \\ \alpha_{5:y} = 68\sqrt{2} \\ \alpha_{6:y} = 715 \\ \quad \text{etc.} \end{cases}$$

Making proper substitution of (30), (32), (33) into (26), we obtain

$$(34)\quad \begin{cases} M_{z_y} = \sigma_x^2 \\ \sigma_{z_y} = \sqrt{\dfrac{2\,(s-r)}{r\,(s-1)}}\,\sigma_x^2 \\ \alpha_{3:z_y} = \dfrac{2\,(s-2r)}{s-2}\sqrt{\dfrac{2\,(s-1)}{r\,(s-r)}} \\ \alpha_{4:z_y} - 3 = \dfrac{12\,(s-1)\,(s^2+s-6rs+6r^2) - 6s\,(r-1)\,(s-r-1)}{r\,(s-r)\,(s-2)\,(s-3)} \end{cases}$$

For an infinite parent population, i.e., $s \to \infty$, we have

$$(35)\quad \begin{cases} M_{z_y} = \sigma_x^2 \\ \sigma_{z_y} = \sqrt{\dfrac{2}{r}}\,\sigma_x^2 \\ \alpha_{3:z_y} = 2\sqrt{\dfrac{2}{r}} \\ \alpha_{4:z_y} - 3 = \dfrac{12}{r} \end{cases}$$

Now again with reference to Pearsonian Types of Curves for which the mode is at $t = a$, we have for the mode of the distribution of sample means z_y,

$$(36)\quad t = \frac{z_y - M_{z_y}}{\sigma_{z_y}} = -\frac{\alpha_{3:z_y}}{2(1+2\delta_{z_y})} \text{ where}$$

$$(37)\quad \begin{aligned} \delta_{z_y} &= \frac{2\alpha_{4:z_y} - 3\alpha_{3:z_y}^2 - 6}{\alpha_{4:z_y} + 3} \\ &= 2 - \frac{(s-3)\,[4(s-2r)^2(s-1) + 2r(s-2)^2(s-r)]}{(s-2)[2(s-1)(s^2+s-6rs+6r^2) + r(s-r)(s-2)(s-3) - s(r-1)(s-r-1)} \end{aligned}$$

Substituting (34) into (36), we obtain

$$(38)\quad \frac{z_y - \sigma_x^2}{\sqrt{\dfrac{2(s-r)}{r(s-1)}}\,\sigma_x^2} = -\frac{s-2r}{s-2}\sqrt{\frac{2(s-1)}{r(s-r)}}\cdot\frac{1}{1+2\delta_{z_y}}$$

By Theorem IV, the best approximated 'most probable value' of the standard deviation of the parent population is obtained from (38) by changing the sign of the right member and replacing z_y by m_2. Thus we have

$$\frac{m_2 - \sigma_x^2}{\sqrt{\dfrac{2(s-r)}{r(s-1)}}\,\sigma_x^2} = \frac{s-2r}{s-2}\sqrt{\frac{2(s-1)}{r(s-r)}}\cdot\frac{1}{1+2\delta_{z_y}}$$

Solving for σ_x, which is now the best approximated 'most probable value' and should therefore be denoted by $\hat{\sigma}_x$, we then have

$$\hat{\sigma}_x^2 = \hat{\mu}_{2:x} = \frac{m_2}{1 + \dfrac{2(s-2r)}{r(s-2)(1+2\delta_{z_y})}} \tag{39}$$

The best approximate 'most probable value' of the standard deviation may therefore be written down as

$$\hat{\sigma}_x = \sigma_s \cdot \frac{1}{\sqrt{1 + \dfrac{2(s-r)}{r(s-2)(1+2\delta_{z_y})}}} \quad \text{where } \sigma_s = \sqrt{m_2}$$

This formula is, of course, subject to a systematic error that arises from the fact that we employ the square root of the best estimated 'most probable value' of the variance. It may be shown, however, that when r is large, the error is small.[14]

Consequently, we have the following theorem:

Theorem VIII. For a normal parent population, the best approximated 'most probable value' of the standard deviation of the parent population is equal to

$$\frac{\sigma_s}{\sqrt{1 + \dfrac{2(s-2r)}{r(s-2)(1+2\delta_{z_y})}}}$$

where σ_s is the standard deviation of an observed sample from the parent population and σ_{z_y} is a function of r and s as expressed in (37).

It is interesting to note from (39) that when $s = 2r$, $\hat{\sigma}_x = \sigma_s$ provided $\delta_{z_y} \neq -\frac{1}{2}$. However, from (37), δ_{z_y} cannot be equal to $-\frac{1}{2}$ in the case of $s = 2r$, where s and r are both positive integers. Consequently, we may state this fact in another theorem:

Theorem IX. If a sample is composed of exactly half of the variates of a normal parent population, then the best approximated 'most probable value' of the standard deviation of that parent population is equal to the standard deviation of an observed sample from it.

For an infinite parent population, (39) yields on reduction

$$\hat{\sigma}_x = \sigma_s \sqrt{\frac{r}{r+2}} \quad \text{for } \sigma_{z_y} = 0 \text{ when } s \to \infty. \tag{40}$$

[14] Professor H. C. Carver has worked out a relation between the most probable value of x^2 and that of x by assuming that the latter is distributed according to a Type III distribution. With his permission, I state the result as follows:

$$\text{M. P. V. } x^p = (\text{M. P. V. } x)^p \left(\frac{\lambda^2 - p}{\lambda^2 - 1}\right)^p$$

where $\lambda = \dfrac{M_x}{\sigma_x}$ and M_x = the distance of the mean from the origin.

Theorem X. For an infinite normal parent population, the best approximated 'most probable value' of the standard deviation of the parent population is equal to the standard deviation of an observed sample multiplied by $\sqrt{\frac{r}{r+2}}$.

Section III. Distribution of the Hypothetical Means of the Parent Population

In the preceding two sections, we have obtained the best approximated 'most probable value' of the mean and the best approximated 'most probable value' of the standard deviation of a parent population assumed to be normal. We are now in the position to characterize the distribution of these hypothetical means by assuming that the best approximated 'most probable value' of the mean of the parent population be its mean and the best approximated 'most probable value' of the standard deviation of the parent population be its standard deviation. Such a characterization is subject to its own probable error.

Due to the fact that our parent population is normal by assumption, formulae (4), which we are to use this time, have to be modified by the proper substitution of the recursion relation of the moments of a normal distribution [See (21)]. After such modifications, they assume the following forms:

$$(41)\quad \begin{cases} M_{z_x} = M_x \\ \bar{\mu}_{2:z_x} = \dfrac{s}{r^2} P_2 \bar{\mu}_{2:x} \\ \bar{\mu}_{3:z_x} = 0 \\ \bar{\mu}_{4:z_x} = \dfrac{3s}{r^4} (P_4 + P_2^2 s)\bar{\mu}_{2:x}^2 \\ \bar{\mu}_{5:z_x} = 0 \\ \bar{\mu}_{6:z_x} = \dfrac{15s}{r^6} (P_6 + 3P_4P_2 s + P_2^3 s^2)\bar{\mu}_{2:x}^3 \end{cases}$$

In accordance with Theorems II and III, we therefore have for the distribution of the means of the parent population the following:

$$(42)\quad \begin{cases} M_{M_x} = m_1 \\ \bar{\mu}_{2:M_x} = \bar{\mu}_{2:z_x} = \dfrac{s}{r^2} P_2 \bar{\mu}_{2:x} = \dfrac{s}{r^2} P_2 \hat{\mu}_{2:x} \\ \bar{\mu}_{3:M_x} = -\bar{\mu}_{3:z_x} = 0 \\ \bar{\mu}_{4:M_x} = \bar{\mu}_{4:z_x} = \dfrac{3s}{r^4} (P_4 + P_2^2 s)\bar{\mu}_{2:x}^2 = \dfrac{3s}{r^4} (P_4 + P_2^2 s)\hat{\mu}_{2:x}^2 \\ \bar{\mu}_{5:M_x} = -\bar{\mu}_{5:z_x} = 0 \\ \bar{\mu}_{6:M_x} = \bar{\mu}_{6:z_x} = \dfrac{15s}{r^6} (P_6 + 3P_4P_2 s + P_2^3 s^2)\hat{\mu}_{2:x}^3 \end{cases}$$

$$= \frac{15s}{r^6} (P_6 + 3P_4P_2 s + P_2^3 s^2)\mu_{2:x}^3$$

Consequently

$$(43)\qquad \begin{cases} M_{M_x} = m_1 \\ \sigma_{M_x} = \sqrt{\dfrac{s-r}{r(s-1)}}\,\hat{\sigma}_x \\ \alpha_{3:M_x} = 0 \\ \alpha_{4:M_x} - 3 = -\dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)} \end{cases}$$

For an infinite parent population, i.e., $s \to \infty$, we have

$$(44)\qquad \begin{cases} M_{M_x} = m_1 \\ \sigma_{M_x} = \dfrac{1}{\sqrt{r}}\hat{\sigma}_x = \dfrac{1}{\sqrt{r}}\sqrt{\dfrac{r}{r+2}}\,\sigma_s = \dfrac{\sigma_s}{\sqrt{r+2}}, \quad \text{[from (40)]} \\ \alpha_{3:M_x} = 0 \\ \alpha_{4:M_x} - 3 = 0 \end{cases}$$

Now if we can find the equation of the curve associated with the distribution of the means of the parent population, we shall be able to ascertain the probability that a mean lies within certain limits after a sample from the parent population has once been observed.

Let us illustrate this by again referring to the same problem of the heights of 1000 freshman students as recorded in Table I. Considering this as our parent population which is almost normal with $s = 1000$, we take every tenth individual height from the original list in which the 1000 heights are tabulated. Thus we obtain a sample with $r = 100$. The frequency distribution of these 100 individual heights is shown in Table IV.

TABLE IV

Sample of 100 Heights Selected from the Parent Population of 1000 from Table I

Class	Frequency
62.5–64.4	9
64.5–66.4	16
66.5–68.4	31
68.5–70.4	29
70.5–72.4	13
72.5–74.4	2

We compute the mean, the standard deviation, the skewness, and the fourth moment about the mean of this sample:

$$m_1 = 67.99$$
$$m_2 = 5.415,2 \qquad \sigma_s = 2.327,058$$
$$m_3 = -1.549,872 \qquad \alpha_{3:s} = -1.229,91$$
$$m_4 = 71.615,158 \qquad \alpha_{4:s} = 2.442,17$$

From Theorem VII,

$$\hat{M}_x = 67.99$$

From (37) and (39), we obtain

$$\delta_{xy} = -.099,833$$
$$\hat{\mu}_{2:x} = 5.328,067$$

Substituting into (42), we have

$$M_{M_x} = 67.99$$
$$\bar{\mu}_{2:M_x} = .048,000,603,6 \qquad \sigma_{M_x} = .219,09$$
$$\bar{\mu}_{3:M_x} = 0 \qquad \alpha_{3:M_x} = 0$$
$$\bar{\mu}_{4:M_x} = .006,898,429 \qquad \alpha_{4:M_x} = 2.994,03$$
$$\bar{\mu}_{5:M_x} = 0 \qquad \alpha_{5:M_x} = 0$$
$$\bar{\mu}_{6:M_x} = .001,649,027 \qquad \alpha_{6:M_x} = 14.910,37$$

The coefficients of Charlier's Type A Function (3) are as follows:

$$\frac{c_3}{3!} = 0$$

$$\frac{c_4}{4!} = -.000,250$$

$$\frac{c_5}{5!} = 0$$

$$\frac{c_6}{6!} = .000,000,1$$

From the values we are justified in assuming that M_x is normally distributed. We may now ask ourselves concerning the probability that the mean of the parent population, M_x, from which this sample is selected, exceeds 68.5 inches.

$$t = \frac{M_x - M_{M_x}}{\sigma_{M_x}} = \frac{68.5 - 67.99}{.21909} = 2.3278$$

$$P = \int_{2.3278}^{\infty} \phi(t)\, dt = .009962$$

Let us now come back to investigation of the general case for the distribution of the hypothetical means of the parent population. Because there is no definite relation between the values of r and s, except $r \leqq s$, and because, by assumption, our parent population is normal, δ_{z_x} is a function of r and s (22); that is

$$\delta_{z_x} = \frac{2s(r-1)(s-r-1)}{s(r-1)(s-r-1) - r(s-r)(s-2)(s-3)}$$

Consequently, it is necessary for us to investigate for different values of δ_{z_x} with respect to various combinations of r and s before we can tell which Type of Pearson's Curves will best fit the distribution of the means of the parent population. Hence, Table V:

TABLE V

Relation of the Values of δ_{z_x} with Various Combinations of r and s

$s = 10r$	$r \geqq 100,$	$\delta_{z_x} \geqq -.0020$
	$r \geqq 50,$	$\delta_{z_x} \geqq -.0040$
	$r \geqq 10,$	$\delta_{z_x} \geqq -.0189$
$s = 5r$	$r \geqq 100,$	$\delta_{z_x} \geqq -.0040$
	$r \geqq 50,$	$\delta_{z_x} \geqq -.0080$
	$r \geqq 10,$	$\delta_{z_x} \geqq -.0397$
$s = 2r$	$r \geqq 100,$	$\delta_{z_x} \geqq -.0101$
	$r \geqq 50,$	$\delta_{z_x} \geqq -.0204$
	$r \geqq 10,$	$\delta_{z_x} \geqq -.1118$

$s = r + 1,\quad r =$ any finite value, $\quad \delta_{z_x} = 0$

$s =$ any finite value, $\quad r = 1 \quad \delta_{z_x} = 0$

$s \to \infty, \quad r =$ any finite value, $\quad \delta_{z_x} = 0.$

From the above table we observe:

1) For an infinite normal parent population, the frequency distribution of the hypothetical means of the parent population is normal, because both $\alpha_{3:M_x}$ and δ_{z_x} are equal to 0 (See Part I, Section III).

2) For any finite, normal parent population, if $r = 1$, the frequency distribution of the hypothetical means of the parent population is normal.

3) For any finite, normal parent population, if a sample $r = s - 1$ is chosen, the frequency distribution of the hypothetical means of the parent population is normal.

4) For any finite, normal parent population, if s is equal to $5r$ or more and at the same time r is at least equal to fifty, the normal curve is a fair approximation for the distribution of the hypothetical means of the parent population.

5) For the other cases in which $|\delta_{z_x}|$ is not negligibly small, we ought to make further investigation.

Now, to carry out further investigation for the cases where $|\delta_{z_x}|$ is not very small, we need only look back to formulae (43), from which we observe that: $\alpha_{4:M_x} - 3 < 0$ for $s \neq r + 1, r \neq 1$, or s does not approach infinity.

Because of the fact that $\alpha_{3:M_x} = 0$ and $\alpha_{4:M_x} < 3$ is the criterion for Type II,[15] we conclude that Type II will be the best fitting curve for the cases mentioned in 5) above. To obtain this Type II curve we proceed as follows:

Let the equation of the curve associated with the distribution of the hypothetical means of the parent population with which we are concerned be $y = P_{M_x}(t)$. Then

$$\frac{1}{y}\frac{dy}{dt} = \frac{a - t}{b_0 + b_1 t + b_2 t^2} = \frac{a - t}{-b_2(t + R)(R - t)}$$

where

$$R = \frac{-b_1 \pm \sqrt{b_1^2 - 4b_0 b_2}}{2b_2}$$

By proper substitution with the formulae in Part I, Section III, we obatin

$$(45)\qquad \begin{aligned} R &= \frac{-\alpha_{3:M_x} \pm \sqrt{\alpha_{3:M_x}^2 - 4\delta_{z_x}(2 + \delta_{z_x})}}{2\delta_{z_x}} \\ &= \pm\sqrt{\frac{-2}{\delta_{z_x}} - 1} \qquad \text{since } \alpha_{3:M_x} = 0 \quad \text{from (44)} \end{aligned}$$

For the same reason $a = \dfrac{-\alpha_{3:M_x}}{2(1 + 2\delta_{z_x})} = 0$; therefore the differential equation may be rewritten as

$$\frac{1}{y}\frac{dy}{dt} = \frac{t}{b_2(R^2 - t^2)},$$

from which we obtain

$$(46)\qquad y = y_0 (R^2 - t^2)^q \text{ where } q = -\frac{1}{2b_2} = -\frac{1 + 2\delta_{z_x}}{\delta_{z_x}}.$$

Imposing the condition that the total area under the curve be equal to unity, we set

$$1 = \int_{-R}^{R} y\,dt = y_0 \int_{-R}^{R} (R^2 - t^2)^q\,dt$$

[15] Elderton, W. P., *op. cit.*, Table VI, opposite p. 46.

Substituting $t = -R + 2R\mu$, we have

$$1 = y_0 \int_0^1 (2R)^{2q+1} \mu^q (1 - \mu)^q \, du$$

$$= y_0 (2R)^{2q+1} \beta(q + 1, q + 1)$$

$$\therefore y_0 = \frac{1}{(2R)^{2q+1}} \cdot \frac{\Gamma(2q + 2)}{\Gamma(q + 1)\,\Gamma(q + 1)}$$

hence

$$(47) \qquad \begin{aligned} y &= \frac{\Gamma(2q + 2)}{(2R)^{2q+1}\,\Gamma(q + 1)\,\Gamma(q + 1)} (R^2 - t^2)^q \\ &= \frac{1}{2^{2q+1}\sqrt{2q + 3}} \cdot \frac{\Gamma(2q + 2)}{\Gamma(q + 1)\,\Gamma(q + 1)} \left(1 - \frac{t^2}{2q + 3}\right)^q, \end{aligned}$$

where q may be expressed in terms of r and s by means of (46) and (22). Thus

$$(48) \quad q = -\frac{1}{\delta_{zx}} - 2 = \frac{r(s - r)(s - 2)(s - 3) - 5s(r - 1)(s - r - 1)}{2s(r - 1)(s - r - 1)}$$

To sum up: In describing the distribution of the hypothetical means of a parent population from which our sample is chosen, we have the following theorems:

Theorem XI. The frequency distribution of the hypothetical means of an infinite, normal parent population is normal.

Theorem XII. The frequency distribution of the hypothetical means of a finite, normal parent population is normal if $r = s - 1$.

Theorem XIII. The frequency distribution of the hypothetical means of a finite, normal parent population is very nearly normal if s is equal to $5r$ or more and r is at least equal to fifty.

Theorem XIV. The frequency distribution of the hypothetical means of a finite, normal parent population is according to Type II for the cases in which $|\delta_{zx}|$ is not negligibly small.

Section IV. Probable Error of the Mean

To measure the fluctuation of a sample mean from the true mean of the parent population, it is customary to use the term "probable error" to denote the expression:

$$(49) \qquad E_M = 0.6745 \frac{\sigma_x}{\sqrt{r}}$$

where σ_x is the standard deviation of the parent population. As the true value of σ_x is not known, it is the common practice to substitute for it the value $\sqrt{\frac{r}{r-1}}\,\sigma_x'$, where σ_x' is the square root of the expected value of the sample second moment.

Therefore (49) is rewritten as

$$E_M = 0.6745 \frac{\sigma'_x}{\sqrt{r-1}} \tag{50}$$

Still, it should be noted, this expression is an approximation. Now from our theory of inverse sampling, as far as a normal parent population is assumed, we have obtained for the probable error of the mean

$$E_M = 0.6745 \frac{\sigma_s}{\sqrt{r+2}} \tag{51}$$

where σ_s is definitely the standard deviation of an observed sample. Although for large r, (50) and (51) do not differ much, yet (51) is obtained directly in terms of the standard deviation of an observed sample.

To illustrate, consider the same sample of the heights of 100 freshman students (See Table IV) as obtained from an infinite parent population. Since the mean is 67.99 and the standard deviation is 2.327058, the probable error of the mean is

$$E_M \doteq 0.6745 \times \frac{2.327058}{\sqrt{102}} = .1554152;$$

that is, $M_x = 67.99 \pm .1554152$, which shows that the chances are even that the true mean of the parent population lies within the range 67.834,584,8 and 68.145,415,2.

Section V. Distribution of the Hypothetical Variances of the Parent Population

Recalling the fact we have stated in Part III, Section II, that the consideration of the distribution of the second moments of samples about the most probable value of the mean is equivalent to the consideration of a distribution of sample means drawn from a parent population $y_2, y_1, y_3, \cdots y_s$, where $y_i = (x_i - m_1)^2$ since in a normal parent population $\hat{M}_x = m_1$ [See (24)] we can write down in perfect analogy with (12) and (14)

$$\begin{aligned} \bar{\mu}_{n:p} &= (-1)^n \bar{\mu}_{n:x_y} \\ M_p &= m_2 \end{aligned} \tag{52}$$

Now

$$\bar{\mu}_{n:p} = \bar{\mu}_{n:M_y} = \bar{\mu}_{n:} \frac{\sum (x_i - m_1)^2}{N} = \bar{\mu}_{n:\mu_2:x}$$

since we have assumed the mean of the parent population to be its most probable value, i.e., m_1. Hence by virtue of (52) and (34), the frequency distribution of the hypothetical variances of the parent population, which is assumed to be normal, is characterized by

$$(53)\quad \begin{aligned} M_{\bar{\mu}_2:x} &= m_2 \\ \sigma_{\bar{\mu}_2x} = \sigma_{z_y} &= \sqrt{\frac{2(s-r)}{r(s-1)}}\,\sigma_x^2 = \sqrt{\frac{2(s-r)}{r(s-1)}}\,\hat{\sigma}_x^2 \end{aligned}$$

since we assume the most probable value of the variance of the parent population to be its variance.

$$\alpha_{3:\bar{\mu}_2:x} = -\alpha_{3:z_y} = -\frac{2(s-2r)}{s-2}\sqrt{\frac{2(s-1)}{r(s-r)}}$$

$$\alpha_{4:\bar{\mu}_2:x} - 3 = \alpha_{4:z_y} - 3 = \frac{12(s-1)(s^2+s-6rs+6r^2) - 6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}$$

For an infinite parent population, i.e., $s \to \infty$, we have

$$(54)\quad \begin{cases} M_{\bar{\mu}_2:x} = m_2 \\ \sigma_{\bar{\mu}_2:x} = \sqrt{\dfrac{2}{r}}\,\hat{\sigma}_x^2 \\ \alpha_{3:\bar{\mu}_2:x} = -2\sqrt{\dfrac{2}{r}} \\ \alpha_{4:\bar{\mu}_2:x} - 3 = \dfrac{12}{r} \end{cases}$$

Now if we can find the equation of the curve associated with the distribution of the hypothetical variances of the parent population, we shall be able to ascertain the probability that a variance lies between certain specified limits after a sample is drawn from the parent population.

For illustration, we will use the same sample of the heights of 100 freshman students (See Table IV) as selected from a parent population of 1000.

We have $s = 1000, \quad r = 100$

$m_1 = 67.99$

$m_2 = 5.4152$, or $\sigma_s = 2.327058$

From (37) we compute

$$\delta_{z_y} = -.0098$$

As $|\delta_{z_y}|$ is negligibly small, we may be justified in considering $\bar{\mu}_{2:x}$ to be distributed according to Type III (Part I, Section III).

It follows from (39) that

$$\hat{\mu}_{2:x} = 5.32$$

We compute the moments of the distribution of the hypothetical variances in accordance with (53). Thus

$$M_{\bar{\mu}_{2:x}} = 5.4152$$
$$\sigma_{\bar{\mu}_{2:x}} = .556$$
$$\alpha_{3:\bar{\mu}_{2:x}} = .239,946$$
$$\alpha_{4:\bar{\mu}_{2:x}} = 3.055,75$$

If we now wish to ascertain the probability that the variance of the parent population lies between $\bar{\mu}_{2:x} = a = 5.5$ and $\bar{\mu}_{2:x} = b = 6.5$, we first convert a, b into standard units such that $t_a = .1525$ and $t_b = 1.9511$ and then evaluate the following integral:[16]

$$P = \frac{\left(\frac{2}{\alpha_3}\right)^{\frac{4}{\alpha_3^2}}}{\Gamma\left(\frac{4}{\alpha_3^2}\right)} e^{\frac{-4}{\alpha_3^2}} \int_{.1525}^{1.9511} \left(\frac{2}{\alpha_3} + t\right)^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3}t} \, dt$$

But this step is now not necessary since we have access to Tables of Pearson's Type III Function.[17] Hence we find from this table our desired probability.

$$P = .39146$$

In the above numerical example, we are justified in using Type III because $|\delta_{zy}|$ is negligibly small. But for the general case, however, we ought to make further investigation concerning the values of δ_{zy}.

TABLE VI

Relation of Values of δ_{zy} with Various Combinations of r and s

$s = 10r$	$r \geq 100$	$\delta_{zy} \geq -.0098$
	$r \geq 50$	$\delta_{zy} \geq -.0194$
	$r \geq 10$	$\delta_{zy} \geq -.0859$
$s = 5r$	$r \geq 100$	$\delta_{zy} \geq -.0200$
	$r \geq 50$	$\delta_{zy} \geq -.0400$
	$r \geq 10$	$\delta_{zy} \geq -.1983$
$s = 2r$	$r \geq 100$	$\delta_{zy} \geq -.0518$
	$r \geq 50$	$\delta_{zy} \geq -.1073$
	$r \geq 10$	$\delta_{zy} \geq -.7642$

$s \to \infty$, r = any finite value, $\delta_{zy} = 0$.

[16] Elderton, P. E., *op. cit.*, p. 90.

[17] Salvosa, L. R., Tables of Pearson's Type III Functions, *Annals of Mathematical Statistics* Vol. I, No. II.

Recalling that δ_{z_y} is a function of r and s such that

$$\delta_{z_y} = 2 - \frac{(s-3)\{4(s-2r)^2(s-1) + 2r(s-2)^2(s-r)\}}{(s-2)\{2(s-1)(s^2+s-6rs+6r^2) + r(s-r)(s-2)(s-3) - s(r-1)(s-r-1)\}}$$

we construct Table VI of δ_{z_y} for different combinations of s and r.

From Table VI we observe the following facts.

1) For an infinite, normal parent population, the distribution of the hypothetical variances of the parent population is according to Type III.

2) For a finite, normal parent population, if s is at least equal to $5r$ and r at least fifty, the distribution of the hypothetical variances of the parent population is very nearly according to Type III.

3) For the other cases in which δ_{z_y} is not small but negative in sign, the distribution of the hypothetical variances of the parent population needs further investigation.

From Part I, Section III, $k = \dfrac{\alpha_3^2}{4\delta(2+\delta)}$; and since we know that δ is always greater than -2, therefore whether k is positive or negative depends upon whether δ is positive or negative.

Now from Table VI we observe that δ_{z_y} seems to be always negative; hence k is negative. In accordance with the criterion for fitting curves, the frequency distribution of the variances of a normal parent population in such cases is according to Type I, which takes the form:[18]

$$(55) \qquad y = \frac{\Gamma_{(m_1+m_2+2)}}{\Gamma_{(m_1+1)}\Gamma_{(m_2+1)}} \cdot \frac{1}{(R_1-R_2)^{m_1+m_2+1}} (t-R_2)^{m_1}(R_1-t)^{m_2}$$

where

$$m_1 = \frac{a-R_2}{b_2(R_2-R_1)}, \qquad m_2 = \frac{a-R_1}{b_2(R_1-R_2)}$$

R_1, R_2 are the positive and negative roots, respectively, of the equation $b_0 + b_1t + b_2t^2 = 0$ and can be expressed in terms of the first four moments:

$$R_1, R_2 = \frac{-\alpha_3 \pm \sqrt{\alpha_3^2 - 4\delta(2+\delta)}}{2\delta}$$

We may sum up the foregoing in the following theorems:

Theorem XV. The frequency distribution of the hypothetical variances of an infinite, normal parent population is according to Type III.

Theorem XVI. The frequency distribution of the hypothetical variances of a finite, normal parent population approximates to Type III Curve if r and s are of such combinations that $|\delta_{z_y}|$ turns out to be negligibly small.

Theorem XVII. The frequency distribution of the hypothetical variances of

[18] Elderton, W. P., *op. cit.*, p. 54.

a finite, normal parent population is according to Type I in case that δ_{z_y} is not very nearly equal to zero and is negative.

Part IV. Inverse Sampling Associated with a Parent Population Distributed According to Pearson's Type III Function

Instead of a normal parent population as we have assumed throughout our discussion in Part III, we shall assume in this part a parent population which is distributed according to Type III. Therefore, besides the distribution of the hypothetical means and that of the hypothetical variances of the parent population, the distribution of the hypothetical third moments will also be considered. We shall carry out our discussion in practically the same way as we have done in Part III.

Section I. Most Probable Value of the Mean of the Parent Population

We have already obtained a general expression for the most probable value of the mean of the parent population:

$$\hat{M}_x = m_1 - \frac{s-2r}{r(s-2)} \frac{\sigma_x \alpha_{3:x}}{2(1+2\delta_{z_x})}$$

where as before

$$\delta_{z_x} = \frac{2\alpha_{4:z_x} - 3\alpha_{3:z_x}^2 - 6}{\alpha_{4:z_x} + 3}$$

But we are now concerned with a parent population which is distributed according to Type III.

Since the recursion relation of the moments of Type III distribution is of the form

$$\alpha_{n+1} = n\left(\alpha_{n-1} + \frac{\alpha_3}{2}\alpha_n\right) \tag{56}$$

$\alpha_4 = 3(1+\gamma)$ where $\gamma = \dfrac{\alpha_3^2}{2}$

$\alpha_5 = 2\alpha_3(5+3\gamma)$

$\alpha_6 = 5(3+13\gamma+6\gamma^2)$

$\alpha_7 = 3\alpha_3(35+77\gamma+30\gamma^2)$

$\alpha_8 = 7(15+170\gamma+261\gamma^2+90\gamma^3)$

$\alpha_9 = 4\alpha_3(315+1652\gamma+2007\gamma^2+630\gamma^3)$

$\alpha_{10} = 9(105+2450\gamma+8435\gamma^2+8658\gamma^3+2520\gamma^4)$

$\alpha_{11} = 5\alpha_3(3456+35266\gamma+91971\gamma^2+82962\gamma^3+22680\gamma^4)$

$\alpha_{12} = 11(945+39375\gamma+252245\gamma^2+537777\gamma^3+437490\gamma^4+113400\gamma^5)$

etc.

it follows from (5) that for a Type III distribution of the parent population

$$(57)\quad \begin{cases} M_{s_x} = M_x \\ \sigma_{s_x} = \sqrt{\dfrac{s-r}{r(s-1)}}\,\sigma_x \\ \alpha_{3:s_x} = \dfrac{s-2r}{s-2}\sqrt{\dfrac{s-1}{r(s-r)}}\,\alpha_{3:x} \\ \alpha_{4:s_x} - 3 = \dfrac{(s-1)(s^2+s-6rs+6r^2)}{r(s-r)(s-2)(s-3)}\cdot\dfrac{\alpha_{3:x}^2}{2} - \dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)} \end{cases}$$

Therefore for the most probable value of the mean of the parent population, we have the same form as (20):

$$\hat{M}_x = m_1 - \frac{s-2r}{r(s-2)}\,\frac{\sigma_x\alpha_{3:x}}{2(1+2\delta_{s_x})};$$

except now instead of (17)

$$(58)\quad \delta_{s_x} = 2 - \frac{3\alpha_{3:s_x}^2 + 12}{\alpha_{4:s_x}+3}$$

$$= 2 - \frac{(s-3)\{2(s-1)(s-2r)^2\alpha_{3:x}^2 + 8r(s-2)^2(s-r)\}}{(s-2)\{(s-1)(s^2+s-6rs+6r^2)\alpha_{3:x}^2 + 4r(s-2)(s-3)(s-r) - 4s(r-1)(s-r-1)\}}$$

We observe that if $\alpha_{3:x} = 0$, this comes back to the case of normal parent population which we have already treated in Part III.

But if $s \to \infty$ while $\alpha_{3:x}$ is finite, then $\delta_{s_x} = 0$. Therefore, for the limiting case, i.e., when the parent population is infinite, we have

$$(59)\quad \hat{M}_x = m_1 - \frac{1}{2r}\,\sigma_x\,\alpha_{3:x}$$

Since σ_x and $\alpha_{3:x}$ are not known, we impose the condition that they assume their best approximated 'most probable values' respectively. Hence, we rewrite (20), (59) in the following forms:

$$(60)\quad \hat{M}_x = m_1 - \frac{s-2r}{r(s-2)}\,\frac{\hat{\sigma}_x\,\hat{\alpha}_{3:x}}{2(1+2\hat{\delta}_{s_x})}$$

where now

$$(60b)\quad \hat{\delta}_{s_x} = 2 - \frac{(s-3)\,\{2(s-1)\,(s-2r)^2\,\hat{\alpha}_{3:x}^2 + 8r(s-2)^2\,(s-r)\}}{(s-2)\{(s-1)\,(s^2+s-6rs+6r^2)\,\hat{\alpha}_{3:x}^2 + 4r(s-2)\,(s-3)\,(s-r) - 4s(r-1)\,(s-r-1)\}}$$

and for the infinite case

$$\hat{M}_x = m_1 - \frac{1}{2r}\hat{\sigma}_x\,\hat{\alpha}_{3:x} \tag{61}$$

So we state our theorem:

Theorem XVIII. For a parent population which is distributed according to Type III, the best approximated 'most probable value' of the mean is the mean of an observed sample from it minus a correction factor which is a function of r, s, $\hat{\sigma}_x$, and $\hat{\alpha}_{3:x}$.

It is also interesting to note that when $s = 2r$ and $\hat{\delta}_{s_x} \neq -\frac{1}{2}$, then $\hat{M} = m_1$.

Section II. Most Probable Value of the Standard Deviation of the Parent Population

We consider, as we have done in Part III, Section II, ${}_sC_r$ possible samples, each consisting of r variates chosen from a parent population s. The second moment of each sample computed about the most probable value of the mean of the parent population may be written as

$$z_1 = \frac{1}{r}\{(x_1 - \hat{M}_x)^2 + (x_2 - \hat{M}_x)^2 + \cdots + (x_r - \hat{M}_x)^2\}$$

$$z_2 = \frac{1}{r}\{(x_2 - \hat{M}_x)^2 + (x_3 - \hat{M}_x)^2 + \cdots + (x_{r+1} - \hat{M}_x)^2\}$$

....

$$z_{\binom{s}{r}} = \frac{1}{r}\{(x_{s-r+1} - \hat{M}_x)^2 + (x_{s-r+2} - \hat{M}_x)^2 + \cdots + (x_s - \hat{M}_x)^2\}$$

If we write $(x_i - \hat{M}_x)^2 = y_i$, the above may be considered as a distribution of sample means drawn from a parent population $y_1, y_2, y_3, \cdots y_s$. Therefore, as (27),

$$\mu_{n:y} = \bar{\mu}_{2n:x} + \binom{2n}{1}\bar{\mu}_{2n-1:x}(M_x - \hat{M}_x) + \binom{2n}{2}\bar{\mu}_{2n-2:x}(M_x - \hat{M}_x)^2 + \binom{2n}{3}\bar{\mu}_{2n-3:x}(M_x - \hat{M}_x)^3 + \cdots + (M_x - \hat{M}_x)^{2n}$$

When we impose the condition that the most probable value of the mean of the parent population be its mean, then the above yields

$$\mu_{n:y} = \bar{\mu}_{2n:x}$$

$$\mu_{1:y} = M_y = \bar{\mu}_{2:x}$$

Consequently

$$\bar{\mu}_{k:y} = \frac{\sum (y - M_y)^k}{N} = \mu_{k:y} - \binom{k}{1} \mu_{k-1:y} M_y + \binom{k}{2} \mu_{k-2:y} M_y^2$$
$$- \cdots + (-1)^k M_y^k$$
$$= \bar{\mu}_{2k:x} - \binom{k}{1} \bar{\mu}_{2k-2:x} \bar{\mu}_{2:x} \dot{-} \binom{k}{2} \bar{\mu}_{2k-4:x} \bar{\mu}_{2:x}^2$$
$$- \cdots + (-1)^k \bar{\mu}_{2:x}^k.$$

Now from the fact that we assume a Type III distribution for our parent population, therefore we have

$$(62)\quad \begin{cases} \bar{\mu}_{2:y} = \bar{\mu}_{4:x} - \bar{\mu}_{2:x}^2 = (3\gamma + 2)\sigma_x^4 \\ \bar{\mu}_{3:y} = \bar{\mu}_{6:x} - 3\bar{\mu}_{4:x}\bar{\mu}_{2:x} + 2\bar{\mu}_{2:x}^3 = (30\gamma^2 + 56\gamma + 8)\sigma_x^6 \\ \bar{\mu}_{4:y} = \bar{\mu}_{8:x} - 4\bar{\mu}_{6:x}\bar{\mu}_{2:x} + 6\bar{\mu}_{4:x}\bar{\mu}_{2:x}^2 - 3\bar{\mu}_{2:x}^4 \\ \quad = (630\gamma^3 + 1707\gamma^2 + 948\gamma + 60)\sigma_x^8 \\ \qquad \text{etc.} \end{cases}$$

Substituting (62) into (26), we have

$$(63)\quad \begin{cases} M_{z_y} = \sigma_x^2 \\ \sigma_{z_y} = \sigma_x^2 \sqrt{\dfrac{s - r}{r(s - 1)}(3\gamma + 2)} \\ \alpha_{3:z_y} = \dfrac{s - 2r}{s - 2} \sqrt{\dfrac{s - 1}{r(s - r)}} \cdot \dfrac{30\gamma^2 + 56\gamma + 8}{(3\gamma + 2)^{3/2}} \\ \alpha_{4:z_y} - 3 = \dfrac{(s - 1)(s^2 + s - 6rs + 6r^2)}{r(s - r)(s - 2)(s - 3)} \cdot \dfrac{630\gamma^3 + 1680\gamma^2 + 912\gamma + 48}{(3\gamma + 2)^2} \\ \qquad - \dfrac{6s(r - 1)(s - r - 1)}{r(s - r)(s - 2)(s - 3)}. \end{cases}$$

For an infinite parent population, the above yields by allowing $s \to \infty$

$$(64)\quad \begin{cases} M_{z_y} = \sigma_x^2 \\ \sigma_{z_y} = \sigma_x^2 \sqrt{\dfrac{3\gamma + 2}{r}} \\ \alpha_{3:z_y} = \dfrac{1}{\sqrt{r}} \dfrac{30\gamma^2 + 56\gamma + 8}{(3\gamma + 2)^{3/2}} \\ \alpha_{4:z_y} - 3 = \dfrac{1}{r} \dfrac{630\gamma^3 + 1680\gamma^2 + 912\gamma + 48}{(3\gamma + 2)^2} \end{cases}$$

In accordance with (38), we write

$$\frac{z_y - \sigma_x^2}{\sigma_x^2 \sqrt{\frac{s-r}{r(s-1)}}(3\gamma + 2)} = \frac{-\frac{s-2r}{s-2}\sqrt{\frac{s-1}{r(s-r)}}\frac{30\gamma^2 + 56\gamma + 8}{(3\gamma+2)^{3/2}}}{2(1 + 2\delta_{zy})} \tag{65}$$

It follows from Theorem IV that for the mode of the standard deviation of the parent population, we have

$$\frac{\frac{\sum (x - \hat{M}_x)^2}{r} - \hat{\sigma}_x^2}{\hat{\sigma}_x^2 \sqrt{\frac{s-r}{r(s-1)}}(3\gamma + 2)} = \frac{\frac{s-2r}{s-2}\sqrt{\frac{s-1}{r(s-r)}}\frac{30\gamma^2 + 56\gamma + 8}{(3\gamma+2)^{\frac{3}{2}}}}{2(1 + 2\delta_{zy})} \tag{66}$$

where

$$\delta_{zy} = \frac{2\alpha_{4:zy} - 3\alpha_{3:zy}^2 - 6}{\alpha_{4:zy} + 3} \tag{67}$$

which is a function of r, s, and $\alpha_{3:x}$.

Assuming the best approximated 'most probable value' of $\alpha_{3:x}$ for $\alpha_{3:x}$ and remembering that

$$\hat{M}_x = m_1 - \frac{s-2r}{r(s-2)} \frac{\hat{\sigma}_x \hat{\alpha}_{3:x}}{2(1 + 2\hat{\delta}_{zx})},$$

we write (66) in the form of

$$m_2 + g^2 \frac{\hat{\sigma}_x^2 \alpha_{3:x}^2}{(1 + 2\hat{\delta}_{zx})^2} - \hat{\sigma}_x^2 = g \frac{30\hat{\gamma}^2 + 56\hat{\gamma} + 8}{(3\hat{\gamma} + 2)(1 + 2\hat{\delta}_{zy})} \hat{\sigma}_x^2$$

$$\hat{\sigma}_x^2 = \frac{m_2}{1 + g\frac{30\hat{\gamma}^2 + 56\hat{\gamma} + 8}{(3\hat{\gamma} + 2)(1 + 2\hat{\delta}_{zy})} - \frac{2g^2\hat{\gamma}}{(1 + 2\hat{\delta}_{zx})^2}} \tag{68}$$

where

$$g = \frac{s-2r}{2r(s-2)},$$

$\hat{\delta}_{zx}$ = (60.b) where $\alpha_{3:x}$ is replaced by $\hat{\alpha}_{3:x}$

$\hat{\delta}_{zy}$ = (67) where $\alpha_{3:x}$ is replaced by $\hat{\alpha}_{3:x}$

$$\gamma = \frac{\hat{\alpha}_{3:x}^2}{2}$$

We rewrite (68) in the abridged form:

$$\hat{\sigma}_x^2 = \frac{m_2}{\phi^2(\hat{\alpha}_{3:x}, r, s)} \tag{69}$$

or

$$\hat{\sigma}_x = \frac{\sigma_s}{\phi(\hat{\alpha}_{3:x}, r, s)}$$

where

$$\phi(\hat{\alpha}_{3:x}, r, s) = \sqrt{1 + g\frac{30\hat{\gamma}^2 + 56\hat{\gamma} + 8}{(3\hat{\gamma} + 2)(1 + 2\hat{\delta}_{z_y})} - \frac{2g^2\hat{\gamma}}{(1 + 2\hat{\delta}_{z_x})^2}}$$

and state our theorem:

Theorem XIX. The best approximated 'most probable value' of the standard deviation of a parent population which is assumed to be distributed according to Type III is equal to the standard deviation of an observed sample of it, multiplied by $\dfrac{1}{\phi(\hat{\alpha}_{3:x}, r, s)}$.

For an infinite parent population, $g = \dfrac{1}{2r}$, $\hat{\delta}_{z_x} = 0$ and

$$\hat{\delta}_{z_y} = \frac{2(630\hat{\gamma}^3 + 1680\hat{\gamma}^2 + 912\hat{\gamma} + 48)(3\hat{\gamma} + 2) - 3(30\hat{\gamma}^2 + 56\hat{\gamma} + 8)^2}{(3\hat{\gamma} + 2)[(630\hat{\gamma}^3 + 1680\hat{\gamma}^2 + 912\hat{\gamma} + 48) - 6r(3\hat{\gamma} + 2)^2]}$$

Theorem XX. The best approximated 'most probable value' of the standard deviation of an infinite parent population which is assumed to be distributed according to Type III is equal to the standard deviation of an observed sample of it, multiplied by $\dfrac{1}{\lim\limits_{s\to\infty} \phi(\hat{\alpha}_{3:x}, r, s)}$.

SECTION III. MOST PROBABLE VALUE OF THE SKEWNESS OF THE PARENT POPULATION

Let us again consider ${}_sC_r$ samples, each consisting of r variates chosen from a parent population s. The third moments of each sample computed about the most probable value of the mean of the parent population may be written as

$$Z_1 = \frac{1}{r}\{(x_1 - \hat{M}_x)^3 + (x_2 - \hat{M}_x)^3 + \cdots + (x_r - \hat{M}_x)^3\}$$

$$Z_2 = \frac{1}{r}\{(x_2 - \hat{M}_x)^3 + (x_3 - \hat{M}_x)^3 + \cdots + (x_{r+1} - \hat{M}_x)^3\} \tag{70}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$Z_{\binom{s}{r}} = \frac{1}{r}\{(x_{s-r+1} - \hat{M}_x)^3 + (x_{s-r+2} - \hat{M}_x)^3 + \cdots + (x_s - \hat{M}_x)^3\}$$

If we write $(x_i - \hat{M}_x)^3 = w_i$, the above may be considered as a distribution of sample means drawn from a parent population $w_1, w_2, w_3, \cdots w_s$. Consequently in accordance with (5), we have

$$(71) \quad \begin{cases} M_{z_w} = M_w \\ \sigma_{z_w} = \sqrt{\dfrac{s-r}{r(s-1)}}\, \sigma_w \\ \alpha_{3:z_w} = \dfrac{s-2r}{s-2} \sqrt{\dfrac{s-1}{r(s-r)}}\, \alpha_{3:w} \\ \alpha_{4:z_w} - 3 = \dfrac{(s-1)(s^2+s-6rs+6r^2)}{r(s-r)(s-2)(s-3)} \{\alpha_{4:w} - 3\} \end{cases}$$
$$- \frac{6s(r-1)(s-r-1)}{r(s-1)(s-2)(s-3)}.$$

Let us write the analogous form of (27):

$$(72) \quad \begin{aligned} \mu_{n:w} &= \frac{1}{N}\sum w^n = \frac{1}{N}\sum (x - M_x + M_x - \hat{M}_x)^{3n} \\ &= \bar{\mu}_{3n:x} + \binom{3n}{1} \bar{\mu}_{3n-1:x}(M_x - \hat{M}_x) \\ &\quad + \binom{3n}{2} \bar{\mu}_{3n-2:x}(M_x - \hat{M}_x)^2 + \cdots + (M_x - \hat{M}_x)^{3n} \end{aligned}$$

Imposing the same condition as before that M_x assumes its most probable value (i.e., $M_x = \hat{M}_x$), then (72) becomes

$$(73) \quad \begin{aligned} \mu_{n:w} &= \bar{\mu}_{3n:x} \\ \mu_{1:w} &= M_\omega = \bar{\mu}_{3:x} \end{aligned}$$

The kth moment of the distribution of w about its mean will then be

$$(74) \quad \begin{aligned} \bar{\mu}_{k:w} &= \frac{\sum (w - M_w)^k}{N} = \mu_{k:w} - \binom{k}{1} \mu_{k-1:w} M_w \\ &\quad + \binom{k}{2} \mu_{k-2:w} M_w^2 - \cdots + (-1)^k M_w^k \\ &= \bar{\mu}_{3k:x} - \binom{k}{1} \bar{\mu}_{3k-3:x} \bar{\mu}_{3:x} + \binom{k}{2} \bar{\mu}_{3k-6:x} \bar{\mu}_{3:x}^2 \\ &\quad - \cdots + (-1)^k \bar{\mu}_{3:x}^k \end{aligned}$$

Since we assume a Type III distribution of the parent population, we have in accordance with the recursion relation (56)

(75)
$$
\begin{aligned}
M_w &= \bar{\mu}_{3:x} = \alpha_{3:x}\,\sigma_x^3 \\
\bar{\mu}_{2:w} &= \bar{\mu}_{6:x} - \bar{\mu}_{3:x}^2 = (15 + 63\gamma + 30\gamma^2)\,\sigma_x^6 \\
\bar{\mu}_{3:w} &= \bar{\mu}_{9:x} - 3\bar{\mu}_{6:x}\,\bar{\mu}_{3:x} + 2\bar{\mu}_{3:x}^3 \\
&= \alpha_{3:x}\,(1215 + 6417\gamma + 7938\gamma^2 + 2520\gamma^3)\,\sigma_x^9 \\
\bar{\mu}_{4:w} &= \bar{\mu}_{12:x} - 4\bar{\mu}_{9:x}\,\bar{\mu}_{3:x} + 6\bar{\mu}_{6:x}\,\bar{\mu}_{3:x}^2 - 3\bar{\mu}_{3:x}^4 \\
&= (10395 + 423225\gamma + 2722599\gamma^2 + 5851683\gamma^3 \\
&\quad + 4792230\gamma^4 + 1247400\gamma^5)\,\sigma_x^{12}
\end{aligned}
$$

Substituting into (71), we have

(76)
$$
\left\{
\begin{aligned}
M_{z_w} &= \bar{\mu}_{3:x} = \alpha_{3:x}\,\sigma_x^3 \\
\sigma_{z_w} &= \sigma_x^3 \sqrt{\frac{s-r}{r(s-1)}\,(30\gamma^2 + 63\gamma + 15)} \\
\alpha_{3:z_w} &= \frac{s-2r}{s-2}\sqrt{\frac{s-1}{r(s-r)}}\;\frac{\alpha_{3:y}\,(1215 + 6417\gamma + 7938\gamma^2 + 2520\gamma^3)}{(15 + 63\gamma + 30\gamma^2)^{3/2}} \\
\alpha_{4:z_w} - 3 &= \frac{(s-1)(s^2 + s - 6rs + 6r^2)}{r(s-r)(s-2)(s-3)} \\
&\quad \frac{9720 + 417555\gamma + 2707992\gamma^2 + 5840343\gamma^3 + 4789530\gamma^4 + 1247400\gamma^5}{(15 + 63\gamma + 30\gamma^2)^2} \\
&\quad - \frac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}.
\end{aligned}
\right.
$$

Allowing $s \to \infty$, we have for an infinite parent population

(77)
$$
\begin{aligned}
M_{z_w} &= \bar{\mu}_{3:x} = \alpha_{3:x}\,\sigma_x^3 \\
\sigma_{z_w} &= \sigma_x^3 \sqrt{\frac{1}{r}\,(30\gamma^2 + 63\gamma + 15)} \\
\alpha_{3:z_w} &= \frac{1}{\sqrt{r}}\;\frac{\alpha_{3:x}\,(1215 + 6417\gamma + 7938\gamma^2 + 2520\gamma^3)}{(15 + 63\gamma + 30\gamma^2)^{3/2}} \\
\alpha_{4:z_w} - 3 &= \frac{1}{r}\cdot\frac{9720 + 417555\gamma + 2707992\gamma^2 + 5840343\gamma^3 + 4789530\gamma^4 + 1247400\gamma^5}{(15 + 63\gamma + 30\gamma^2)^2}
\end{aligned}
$$

The best approximated 'most probable value' of $\bar{\mu}_{3:x}$ may now be written after the same fashion as in the preceding cases:

$$\hat{\mu}_{3:x} = \frac{\sum (x - \hat{M}_x)^3}{r} - \frac{\sigma_{z_w} \cdot \alpha_{3:z_w}}{2(1 + 2\delta_{z_w})} \tag{78}$$

where

$$\delta_{z_w} = \frac{2\alpha_{4:z_w} - 3\alpha_{3:z_w}^2 - 6}{\alpha_{4:z_w} + 3}$$

Since

$$\frac{\sum (x - \hat{M}_x)^3}{r} = \frac{1}{r} \sum \left(x - m_1 + g \frac{\hat{\sigma}_x \hat{\alpha}_{3:x}}{1 + 2\hat{\delta}_{z_x}} \right)^3 \qquad \text{[from (60)]},$$

and since we assume the best approximated 'most probable values' of the standard deviation and the skewness for the standard deviation and the skewness of the parent population respectively, we obtain from (78)

$$\hat{\sigma}_x^3 \hat{\alpha}_{3:x} = m_3 + 3m_2 g \frac{\hat{\sigma}_x \hat{\alpha}_{3:x}}{1 + 2\hat{\delta}_{z_x}} + g^3 \frac{\hat{\sigma}_x^3 \hat{\alpha}_{3:x}^3}{(1 + 2\hat{\delta}_{z_x})^3}$$
$$- g \frac{\hat{\alpha}_{3:x} (1215 + 6417\hat{\gamma} + 7938\hat{\gamma}^2 + 2520\hat{\gamma}^3)}{(1 + 2\hat{\delta}_{z_w}) (15 + 63\hat{\gamma} + 30\hat{\gamma}^2)} \hat{\sigma}_x^3$$

The change of $\hat{\mu}_{3:x}$ to $\hat{\sigma}_x^3 \hat{\alpha}_{3:x}$ involves a systematic error although it is small.
Again by proper substitution of (69) we have

$$\frac{\sigma_s^3 \hat{\alpha}_{3:x}}{\phi^3(\hat{\alpha}_{3:x}, r, s)} = \sigma_s^3 \alpha_{3:s} + 3 \sigma_s^3 g \frac{\hat{\alpha}_{3:x}}{\phi (\hat{\alpha}_{3:x}, r, s) (1 + 2\hat{\delta}_{z_x})}$$
$$+ g^3 \frac{\sigma_s^3 \hat{\alpha}_{3:x}^3}{\phi^3 (\hat{\alpha}_{3:x}, r, s) (1 + 2\hat{\delta}_{z_x})^3}$$
$$- g \frac{\hat{\alpha}_{3:x} \sigma_s^3 (1215 + 6417\hat{\gamma} + 7938\hat{\gamma}^2 + 2520\hat{\gamma}^3)}{\phi^3 (\hat{\alpha}_{3:x}, r, s) \cdot (15 + 63\hat{\gamma} + 30\hat{\gamma}^2) (1 + 2\hat{\delta}_{z_w})}.$$

Solving for $\alpha_{3:s}$, we have

$$\alpha_{3:s} = \frac{\hat{\alpha}_{3:x}}{\phi^3 (\hat{\alpha}_{3:x}, r, s)} \left[1 - \frac{3g \phi^2 (\hat{\alpha}_{3:x}, r, s)}{(1 + 2\hat{\delta}_{z_x})} - \frac{2\hat{\gamma} g^3}{(1 + 2\hat{\delta}_{z_x})^3} \right.$$
$$\left. + \frac{g (1215 + 6417\hat{\gamma} + 7938\hat{\gamma}^2 + 2520\hat{\gamma}^3)}{(1 + 2\hat{\delta}_{z_w}) (15 + 63\hat{\gamma} + 30\hat{\gamma}^2)} \right]. \tag{79}$$

Since the right member of (79) is a function of $\hat{\alpha}_{3:x}$, r, and s, therefore the most probable value of $\alpha_{3:x}$ may be approximated when we are given s, r, and the skewness of an observed sample. As it is an algebraic equation of high order in $\hat{\alpha}_{3:x}$ and is so much involved, even approximation presents practical difficulty. However, if once $\hat{\alpha}_{3:x}$ is approximated, $\hat{\sigma}_x$ and $\hat{M}_x$ can be easily obtained from (60) and (68).

Theorem XXI. For the best approximated 'most probable value' of the skewness of a parent population which is assumed to be distributed according to Type III, we must approximate it from equation (79), in which the skewness of an observed sample is expressed as a function of s, r, and the best approximated 'most probable value' of the skewness of the parent population.

To construct a table for the best approximated 'most probable value' $\hat{\alpha}_{3:x}$ corresponding to $\alpha_{3:s}$ for particular values of r, s, we should first reverse the process by assigning different values of $\hat{\alpha}_{3:x}$ so as to obtain $\alpha_{3:s}$; then by the way of interpolation, we shall be able to obtain $\hat{\alpha}_{3:x}$ for a particular $\alpha_{3:s}$.

TABLE VII

Relation of the Sample Skewness and the Best Approximated 'Most Probable Value' of the Parent Population Whose Distribution is According to Type III

$(s \to \infty, \quad r = 100)$

$\alpha_{3:s}$	$\hat{\alpha}_{3:x}$
.1	.0784
.2	.1568
.3	.2373
.4	.3164
.5	.3969
.6	.4776
.7	.5589
.8	.6410
.9	.7239
1.0	.8072
1.1	.8905
1.2	.9737
1.3	1.0567
1.4	1.1392
1.5	1.2211
1.6	1.3022
1.7	1.3791
1.8	1.4578
1.9	1.5355
2.0	1.6122
2.1	1.6828
2.2	1.7609
2.3	1.8303
2.4	1.9024
2.5	1.9670
2.6	2.0371

For $s \to \infty$ and $r = 100$, we have computed the best approximated 'most probable value' of $\alpha_{3:x}$ corresponding to the values of $\alpha_{3:s}$ from .1 to 2.6 as shown in Table VII.

The computation for such a table is laborious because it involves the computation of $\hat{\delta}_{z_x}$, $\hat{\delta}_{z_y}$, and $\hat{\delta}_{z_w}$ which are in turn functions of $\hat{\alpha}_{3:z_x}$ and $\hat{\alpha}_{4:z_x}$, $\hat{\alpha}_{3:z_y}$ and $\hat{\alpha}_{4:z_y}$, and $\hat{\alpha}_{3:z_w}$ and $\hat{\alpha}_{4:z_w}$, respectively.

Section IV. Distribution of the Hypothetical Means of the Parent Population

Since we have obtained in the preceding sections expressions for the best approximated 'most probable values' of the mean, the standard deviation and the skewness of a parent population which is assumed to be distributed according to Type III, we are now in the position to characterize the distribution of the hypothetical means of the parent population with the assumption that the best approximated 'most probable values' of the mean, the standard deviation, and the skewness be the mean, the standard deviation, and the skewness of the parent population.

Basing upon the fundamental relations in (15), we write down the characteristics of the distribution of the hypothetical means of the parent population as follows:

$$(80)\quad \begin{cases} M_{M_x} = m_1 \\ \sigma_{M_x} = \sigma_{z_x} = \hat{\sigma}_x \sqrt{\dfrac{s-r}{r(s-1)}} = \dfrac{\sigma_s}{\phi(\hat{\alpha}_{3:x}, s, r)} \sqrt{\dfrac{s-1}{r(s-1)}} \\ \alpha_{3:M_x} = -\alpha_{3:z_x} = -\dfrac{s-2r}{s-2} \sqrt{\dfrac{s-1}{r(s-r)}}\, \hat{\alpha}_{3:x} \\ \alpha_{4:M_x} - 3 = \alpha_{4:z_x} - 3 \\ \qquad = \dfrac{(s-1)(s^2+s-6rs+6r^2)}{r(s-r)(s-2)(s-3)} \left[\dfrac{3\hat{\alpha}_{3:x}^2}{2}\right] - \dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}. \end{cases}$$

where $\phi(\hat{\alpha}_{3:x}, s, r)$ is given in (69).

For an infinite parent population by allowing $s \to \infty$, we obtain from the above:

$$(81)\quad \begin{cases} M_{M_x} = m_1 \\ \sigma_{M_x} = \dfrac{1}{\sqrt{r}} \dfrac{\sigma_s}{\phi(\hat{\alpha}_{3:x}, r)} \\ \alpha_{3:M_x} = \dfrac{1}{\sqrt{r}} \hat{\alpha}_{3:x} \\ \alpha_{4:M_x} - 3 = \dfrac{3}{2r} \hat{\alpha}_{3:x}^2 \end{cases}$$

where $\phi(\hat{\alpha}_{3:x}, r) = \lim_{s \to \infty} \phi(\hat{\alpha}_{3:x}, s, r)$

Since we observe that the moments of the distribution of the hypothetical means are expressed in terms of $\hat{\alpha}_{3:x}$, it is therefore necessary for us to find the best approximated 'most probable value' of the skewness of a parent population before we attempt to obtain the frequency function associated with the distribution of these hypothetical means.

Numerical illustration. A sample of 100 weights of freshman students is observed and the frequency distribution is given in Table VIII.

TABLE VIII

Weights of 100 Freshman Students

(Original Measurements Correct to Nearest Pound)

Class Mark	Frequency
109.5	4
119.5	11
129.5	25
139.5	34
149.5	14
159.5	8
169.5	0
179.5	3
189.5	1
	100

The first four moments are computed

$$m_1 = 138.3$$
$$\sigma_s = 14.6366$$
$$\alpha_{3:s} = .81099$$
$$\alpha_{4:s} = 4.47644$$

Now, assuming this sample is drawn from an infinite parent population which is assumed to be distributed according to Type III, we wish to find (a) the best approximated 'most probable values' of the mean, the standard deviation, and the skewness of the parent population, and (b) the probability that the mean of the parent population lies between $M_x = 135$ and $M_x = 140$.

By interpolation from Table VII, we obtain the best approximated 'most probable value' of the skewness of the parent population:

$$\hat{\alpha}_{3:x} = .6501$$

From (69) and (61) we obtain

$$\hat{\sigma}_x = 14.5452$$
$$\hat{M}_x = 138.25272, \qquad \phi(\hat{\alpha}_{3:x}, r) = 1.006279$$

From (81) we have

$$M_{M_x} = 138.3$$
$$\sigma_{M_x} = 1.45452$$
$$\alpha_{3:M_x} = .06501$$
$$\alpha_{4:M_x} = 3.00633945$$

$\delta_{z_x} = 0$, the distribution of M_x is associated with Type III Function; hence for the probability that M_x lies between $M_x = 135$ and $M_x = 140$, we again refer to Tables of Pearson's Type III Function prepared by L. R. Salvosa,[19] and we obtain in this case

$$P = .8677592$$

Since the determination of the best fit of a frequency curve in general depends upon the values of α_3, α_4, and k, and since in the present case each of them is a function of s, r, and $\alpha_{3:x}$, we are therefore not able to tell the type of curve to be used until we know s, r, and $\hat{\alpha}_{3:x}$.

For the infinite case, however, as we have illustrated Type III Function may always be used because

$$\delta_{z_x} = \frac{2\alpha_{4:z_x} - 3\alpha^2_{3:z_x} - 6}{\alpha_{4:z_x} + 3} = \frac{2\alpha_{4:M_x} - 3\alpha^2_{3:M_x} - 6}{\alpha_{4:M_x} + 3} = 0$$

holds for all values of $\hat{\alpha}_{3:x}$ and r. We therefore conclude that the hypothetical means of an infinite parent population which is itself distributed according to Type III is distributed according to Type III. Hence

Theorem XXII. The hypothetical means of an infinite parent population is distributed according to Type III if the parent population is assumed to be distributed according to Type III.

Section V. Distribution of the Hypothetical Variances of the Parent Population

Parallel to Part III, Section V, the distribution of the hypothetical variances of a parent population which is assumed to be distributed according to Type III can be described. The fundamental relation of Theorems II and III hold:

$$\bar{\mu}_{2n:p} = \bar{\mu}_{2n:z_y} \qquad \text{or} \qquad \alpha_{2n:p} = \alpha_{2n:z_y}$$
$$\bar{\mu}_{2n+1:p} = -\bar{\mu}_{2n+1:z_y} \qquad , \qquad \alpha_{2n+1:p} = -\alpha_{2n+1:z_y}$$

But now $M_p = \dfrac{\sum (x - \hat{M}_x)^2}{r}$ (See Part IV, Section II)

$$M_p = \frac{1}{r}\sum\left(x - m_1 + g\frac{\hat{\sigma}_x\hat{\alpha}_{3:x}}{1 + 2\hat{\delta}_{z_x}}\right)^2 \tag{82}$$

$$= m_2 + g^2\frac{\sigma^2_S\,\hat{\alpha}^2_{3:x}}{(1 + 2\hat{\delta}_{z_x})^2\phi^2(\hat{\alpha}_{3:x}, r, s)} \quad \text{[from (60)]}.$$

[19] Salvosa, L. R., *Annals of Mathematical Statistics* Vol. I, No. II, 1930..

Upon the same assumption that the best approximated 'most probable values' of the mean, the standard deviation and the skewness be the mean, the standard deviation, and the skewness of the parent population, the distribution of $\bar{\mu}_{2:x}$ is characterized by

$$
(83)\begin{cases}
M_{\bar{\mu}_{2:x}} = m_2 + g^2 \dfrac{\sigma_S^2 \hat{\alpha}_{3:x}^2}{(1 + 2\hat{\delta}_{zx})^2 \phi^2(\hat{\alpha}_{3:x}, r, s)} = m_2 \left\{ 1 + \dfrac{g^2 \hat{\alpha}_{3:x}^2}{(1 + 2\hat{\delta}_{zx})^2 \phi^2(\hat{\alpha}_{3:x}, r, s)} \right\} \\
\sigma_{\bar{\mu}_{2:x}} = \sigma_{zy} = \sqrt{\dfrac{s - r}{r(s-1)}}\, \sigma_y = \sqrt{\dfrac{s-r}{r(s-1)}(3\gamma + 2)}\, \sigma_x^2 \\
\qquad = \sqrt{\dfrac{s-r}{r(s-1)}(3\hat{\gamma} + 2)}\, \dfrac{\sigma_S^2}{\phi^2(\hat{\alpha}_{3:x}, r, s)} \\
\alpha_{3:\bar{\mu}_{2:x}} = -\alpha_{3:zy} = -\dfrac{s - 2r}{s - 2} \sqrt{\dfrac{s-1}{r(s-\gamma)}}\, \alpha_{3:y} \\
\qquad = -\dfrac{s-2r}{s-2} \sqrt{\dfrac{s-1}{r(s-r)}} \cdot \dfrac{30\hat{\gamma}^2 + 56\hat{\gamma} + 8}{(3\hat{\gamma} + 2)^{\frac{3}{2}}} \\
\alpha_{4:\bar{\mu}_{2:x}} - 3 = \alpha_{4:zy} - 3 = \dfrac{(s-1)(s^2 + s - 6rs) + 6r^2)}{r(s-r)(s-2)(s-3)} [\alpha_{4:y} - 3] \\
\qquad - \dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)} \\
\qquad = \dfrac{(s-1)(s^2 + s - 6rs + 6r^2)}{r(s-r)(s-2)(s-3)} \left[\dfrac{630\hat{\gamma}^3 + 1680\hat{\gamma}^2 + 912\hat{\gamma} + 48}{(3\hat{\gamma} + 2)^2} \right] \\
\qquad - \dfrac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}.
\end{cases}
$$

For an infinite parent population, we have

$$
(84)\begin{cases}
M_{\bar{\mu}_{2:x}} = m_2 \left\{ 1 + \dfrac{1}{4r^2} \cdot \dfrac{\hat{\alpha}_{3:x}^2}{\phi^2(\hat{\alpha}_{3:x}, r)} \right\} \\
\sigma_{\bar{\mu}_{2:x}} = \sqrt{\dfrac{(3\hat{\gamma} + 2)}{r}} \cdot \dfrac{m_2}{\phi^2(\hat{\alpha}_{3:x}, r)} \\
\alpha_{3:\bar{\mu}_{2:x}} = -\dfrac{1}{\sqrt{r}} \dfrac{30\hat{\gamma}^2 + 56\hat{\gamma} + 8}{(3\hat{\gamma} + 2)^{3/2}} \\
\alpha_{4:\bar{\mu}_{2:x}} - 3 = \dfrac{1}{r} \left\{ \dfrac{630\hat{\gamma}^3 + 1680\hat{\gamma}^2 + 912\hat{\gamma} + 48}{(3\hat{\gamma} + 2)^2} \right\}
\end{cases}
$$

Numerical illustration. Using the same sample in Table VIII, we wish to ascertain the probability that the variance of the parent population lies between

306.25 and 342.25. From $m_1 = 138.3$, $\sigma_s = 14.6366$, $\alpha_{3:s} = .81099$, and $\alpha_{4:s} = 4.47644$, we find from (84)

$$M_{\bar{\mu}_2:x} = 214.232,235$$
$$\sigma_{\bar{\mu}_2:x} = 34.335,74$$
$$\alpha_{3:\bar{\mu}_2:x} = -.495,311$$
$$\alpha_{4:\bar{\mu}_2:x} = 3.463,675,7$$
$$\delta_{z_y} = .105,515,6$$

From Part I, Section III,

$$k = \frac{\alpha^2_{3:\bar{\mu}_2:x}}{4\delta_{z_y}(2 + \delta_{z_y})} = .276 < 1$$

Therefore, the best fitting curve will be Type IV which assumes the form[20]

$$y = y_0 (1 + x^2)^{-m} e^{-\lambda \tan^{-1} x} \quad (85)$$

where

$$x = \frac{t + p}{q},$$

t being in standard units

$$p = \frac{b_1}{2b_2} = \frac{\alpha_3}{2\delta}$$

$$q^2 = \frac{4b_0b_2 - b_1^2}{4b_2^2} = \frac{4\delta(2 + \delta) - \alpha_3^2}{4\delta^2}$$

$$m = \frac{1}{2b_2} = \frac{1 + 2\delta}{\delta}$$

$$\lambda = -\frac{a + p}{b_2 q}$$

$$y_0 = \frac{e^{\frac{\lambda\pi}{2}}}{g(2m - 2, \lambda)} = \frac{1}{F(2m - 2, \lambda)}$$

y_0 is found from Pearson's *Tables for Statisticians and Biometricians*[21] to be 049662.

---

[20] Elderton, W. P., *op. cit.*, p. 64.

[21] Pearson, K., *Tables for Statisticians and Biometricians*, Vol. I, pp. 126–142.

Now the given limits 306.25 and 342.25 of the variance, when expressed in standard units, are

$$t_a = 2.679{,}941$$

$$t_b = 3.728{,}410$$

Therefore the probability that $\bar{\mu}_{2:x}$ lies between $\bar{\mu}_{2:x} = 306.25$ and $\bar{\mu}_{2:x} = 342.25$ is

$$P = y_0 \int_{t_a=2.679,041}^{t_b=3.728,410} (1 + x^2)^{-m} e^{-\lambda \tan^{-1} x} \, dx$$

we find

$$m = 11.477{,}271$$

$$\lambda = 12.940{,}307$$

$$P = .049662 \int_{.08757}^{.36343} (1 + x^2)^{-11.477271} e^{-12.940207 \tan^{-1} x} \, dx$$

By means of Maclaurin-Euler's Interpolation Formula, P is found to be equal to .000,904.

No definite law can be ascertained before we know $\hat{\alpha}_{3:x}$ because, as we have seen, $\alpha_{3:\bar{\mu}_2:x}$ and $\alpha_{4:\bar{\mu}_2:x}$ are both expressed in terms of s, r, and $\hat{\alpha}_{3:x}$. We do not know the value of k, which is a determining factor of the best fitting curve and a function of s, r, $\alpha_{3:\bar{\mu}_2:x}$ and $\alpha_{4:\bar{\mu}_2:x}$, until we know the values of s, r, and $\hat{\alpha}_{3:x}$.

Section VI. Distribution of the Hypothetical Third Moments of the Parent Population About Its Mean

Recalling the fact that the distribution of the third moments of sample means about the most probable value of the mean of the parent population is equivalent to the consideration of a distribution of sample means drawn from a parent population, $w_1, w_2, w_3, \cdots w_s$ where $w_i = (x_i - \hat{M}_x)^3$, so we can write down in accordance with the fundamental relations stated in Theorems II and III:

$$\bar{\mu}_{2n:p} = \bar{\mu}_{2n:z_w} \qquad \alpha_{2n:p} = \alpha_{2n:z_w}$$

or

$$\bar{\mu}_{2n+1:p} = -\bar{\mu}_{2n+1:z_w} \qquad \alpha_{2n+1:p} = -\alpha_{2n+1:z_w}$$

But here $M_p = \dfrac{\sum (x - \hat{M}_x)^3}{r}$; and by the substitution of (60), we have

$$
\begin{aligned}
M_p &= \frac{1}{r} \sum \left(x - m_1 + \frac{g\,\hat{\sigma}_x\,\hat{\alpha}_{3:x}}{1 + 2\hat{\delta}_{z_x}} \right)^3 \\
&= m_3 + 3m_2 \frac{g\,\sigma_s\,\hat{\alpha}_{3:x}}{(1 + 2\hat{\delta}_{z_x})\,\phi(\hat{\alpha}_{3:x}, r, s)} + \frac{g^3\,\sigma_s^3\,\hat{\alpha}_{3:x}^3}{(1 + 2\hat{\delta}_{z_x})^3\,\phi^3(\hat{\alpha}_{3:x}, r, s)}
\end{aligned} \tag{86}
$$

Consequently, with the same assumption that the best approximated 'most probable values' of the mean, the standard deviation, and the skewness be the mean, the standard deviation and the skewness of the parent population, the distribution of $\bar{\mu}_{3:x}$ is characterized by

$$M_{\bar{\mu}_{3:x}} = m_3 + 3m_2 \frac{g\sigma_s \hat{\alpha}_{3:x}}{(1 + 2\hat{\delta}_{z_x})\phi(\hat{\alpha}_{3:x}, r, s)} + \frac{g^3 \sigma_S^3 \hat{\alpha}_{3:x}^3}{(1 + 2\hat{\delta}_{z_x})^3 \phi^3(\hat{\alpha}_{3:x}, r, s)}$$

$$\sigma_{\bar{\mu}_{3:x}} = \sqrt{\frac{s-r}{r(s-1)}}\, \sigma_{z_w} = \sqrt{\frac{s-r}{r(s-1)}}\, (15 + 63\hat{\gamma} + 30\hat{\gamma}^2)\, \hat{\sigma}_x^3$$

$$= \sqrt{\frac{s-r}{r(s-1)}}\, (15 + 63\hat{\gamma} + 30\hat{\gamma}^2) \frac{\sigma_S^3}{\phi^3(\hat{\alpha}_{3:x}, r, S)}$$

$$\alpha_{3:\bar{\mu}_{3:x}} = -\frac{s-2r}{s-2}\sqrt{\frac{s-1}{r(s-r)}}\, \alpha_{3:z_w}$$

(87)

$$= -\frac{s-2r}{s-2}\sqrt{\frac{s-1}{r(s-r)}}\, \frac{\hat{\alpha}_{3:x}(1215 + 6417\hat{\gamma} + 7938\hat{\gamma}^2 + 2520\hat{\gamma}^3)}{(15 + 63\hat{\gamma} + 30\hat{\gamma}^2)^{\frac{3}{2}}}$$

$$\alpha_{4:\bar{\mu}_{3:x}} - 3 = \frac{(s-1)(s^2 + s - 6rs + 6r^2)}{r(s-r)(s-2)(s-3)}[\alpha_{4:z_w} - 3] - \frac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}$$

$$= \frac{(s-1)(s^2 + s - 6rs + 6r^2)}{r(s-r)(s-2)(s-3)} \cdot \frac{9720 + 417555\hat{\gamma} + 2707992\hat{\gamma}^2 + 5840343\hat{\gamma}^3 + 4789530\hat{\gamma}^4 + 1247400\hat{\gamma}^5}{(15 + 63\hat{\gamma} + 30\hat{\gamma}^2)^2}$$

$$- \frac{6s(r-1)(s-r-1)}{r(s-r)(s-2)(s-3)}.$$

For an infinite parent population, we have

(88)

$$\begin{cases} M_{\bar{\mu}_{3:x}} = m_3 + 3m_2 \dfrac{\sigma_S \hat{\alpha}_{3:x}}{2r\phi(\hat{\alpha}_{3:x}, r)} + \dfrac{1}{8r^3} \dfrac{\sigma_S^3 \hat{\alpha}_{3:x}^3}{\phi^3(\hat{\alpha}_{3:x}, r)} \\ \sigma_{\bar{\mu}_{3:x}} = \sqrt{\dfrac{1}{r}}\, (15 + 63\hat{\gamma} + 30\hat{\gamma}^2) \dfrac{\sigma_S^3}{\phi^3(\hat{\alpha}_{3:x}, r)} \\ \alpha_{3:\bar{\mu}_{3:x}} = -\sqrt{\dfrac{1}{r}}\, \dfrac{\hat{\alpha}_{3:x}(1215 + 6417\hat{\gamma} + 7938\hat{\gamma}^2 + 2520\hat{\gamma}^3)}{(15 + 63\hat{\gamma} + 30\hat{\gamma}^2)^{\frac{3}{2}}} \\ \alpha_{4:\bar{\mu}_{3:x}} - 3 = \dfrac{1}{r} \dfrac{9720 + 417555\hat{\gamma} + 2707992\hat{\gamma}^2 + 5840343\hat{\gamma}^3 + 4789530\hat{\gamma}^4 + 1247400\hat{\gamma}^5}{(15 + 63\hat{\gamma} + 30\hat{\gamma}^2)^2} \end{cases}$$

Numerical illustration. Using the same sample in Table VIII, we wish to ascertain the probability that the third moment of the parent population about the mean lies between $\bar{\mu}_{3:x} = 3000$ and $\bar{\mu}_{3:x} = 4000$, still assuming an infinite parent population from which the sample is drawn

$$\hat{\alpha}_{3:x} = .6501$$
$$\phi(\hat{\alpha}_{3:x}, r) = 1.006,279$$

We find from (88)

$$\begin{aligned} M_{\bar{\mu}_{3:x}} &= 2558.137,096 \\ \sigma_{\bar{\mu}_{3:x}} &= 1675.696,37 \\ \alpha_{3:\bar{\mu}_{3:x}} &= -1.187,409,9 \\ \alpha_{4:\bar{\mu}_{3:x}} &= 6.127,551,6 \\ \delta_{zw} &= 0.221,886 \\ k &= 0.714,972 < 1 \end{aligned}$$

Therefore the best fitting curve is Type IV.

From Pearson's *Tables for Statisticians and Biometricians*, Vol. I,[22] we compute

$$y_0 = .000,058,032,3$$

The given limits 3000 and 4000 when expressed in standard units are $t = .263,689$ and $t = .860,455$ respectively. Therefore the probability that $\bar{\mu}_{3:x}$ lies between 3000 and 4000 may be expressed by

$$P = y_0 \int_{t=.263689}^{t=.860455} (1 + x^2)^{-6.506819} e^{-17.443447 \tan^{-1} x} \, dx$$

By means of Maclaurin-Euler's Interpolation Formula, the answer is found to be .267,408,631.

We make the same remark here as we have made in the preceding two sections. That is, since $\alpha_{3:\bar{\mu}_{3:x}}$ and $\alpha_{4:\bar{\mu}_{3:x}}$ are both in terms of s, r and $\hat{\alpha}_{3:x}$,, we cannot determine the value of k which is a function of $\alpha_{3:\bar{\mu}_{3:x}}$ and $\alpha_{4:\bar{\mu}_{3:x}}$ until we know the values of s, r, and $\hat{\alpha}_{3:x}$. Consequently, the curve associated with the distribution of the hypothetical third moments of a parent population of Type III distribution is not known until we know s, r, and $\hat{\alpha}_{3:x}$.

[22] Pearson, K., *op. cit.*, pp. 126–142.