Josef Machek
On a two-sample procedure for testing Student's hypothesis using mean range

Aplikace matematiky, Vol. 4 (1959), No. 3, 211--223

Persistent URL: http://dml.cz/dmlcz/102662

Terms of use:
© Institute of Mathematics AS CR, 1959

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these Terms of use.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project DML-CZ: The Czech Digital Mathematics Library http://dml.cz
ON A TWO-SAMPLE PROCEDURE FOR TESTING STUDENT'S HYPOTHESIS USING MEAN RANGE

JOSEF MACHEK
(Received September 9th 1958)

Stein's two-sample procedure for testing a hypothesis concerning the mean of a normal distribution is modified for the case where the usual mean-square estimate of the standard deviation is replaced by an estimate based on sample range or on mean range of several samples. Expressions for the mean number of necessary observations are derived and some illustrative tables given. Tables of 5% and 10% critical points of the corresponding test statistic are given.

1. INTRODUCTION AND STATEMENT OF PROBLEM

Throughout the paper we shall be concerned with one-sided tests of a hypothesis about the mean of a normal distribution where the value of the population standard deviation is not known and it is nevertheless desired that the power-function of the test pass through two prescribed points. Thus we shall consider tests which have to provide a given protection not only against an unjust rejection of the hypothesis when the mean takes the value specified by it, but also against accepting it when the mean has a given value different from that specified by the hypothesis. Moreover, we shall try to make the test rather simple in application.

The need for such tests is most likely to occur in problems of sampling inspection. An example is provided by acceptance sampling of materials; suppose, e.g., that a lot of material, say a chemical, is to be sampled. A quantitative characteristic such as concentration or percentage content of some ingredient is relevant for the quality of the material. Lots with values of this characteristic smaller than a given standard are perfectly suitable and should be accepted on the average at least in a prescribed proportion of cases, say 95 per cent. With increasing values of the characteristic the quality of lots decreases so that lots in which the relevant characteristic has a value as high as another given constant or higher are unacceptable (we may well imagine
that the use of them is associated with serious difficulties) and the consumer is prepared to run only a risk of no more than one in ten, say, of accepting such lots. If the standard deviation of the results of corresponding analyses were known and some further assumptions (concerning the homogeneity of the lots and the method of sampling) valid, the construction of a sampling plan would present no difficulties of principle. If the standard deviation cannot be regarded as known, the solution is more complicated.

The problem may be described as follows: Let us have a population whose distribution is normal, with mean \( \mu \) and standard deviation \( \sigma \). Both \( \mu \) and \( \sigma \) are unknown. Two numbers are given, \( \mu_0 \) and \( \Lambda \) (\( \Lambda > 0 \)). It is desired to test the hypothesis \( \mu \leq \mu_0 \) against the alternative \( \mu > \mu_0 \) and the test has to satisfy the following conditions (\( P(\mu) \) denotes the power of the test, i.e., the probability that the hypothesis will be rejected when \( \mu \) is the true value of the mean):

\[
\begin{align*}
P(\mu_0) &= \alpha, \\
P(\mu_0 + \Lambda) &= 1 - \beta.
\end{align*}
\]

Here \( \alpha \) and \( \beta \) are previously fixed numbers from the interval \((0, 1)\), usually equal to 0.10, 0.01 or the like. They represent, respectively, the highest probability of rejecting the hypothesis if it is in fact true and the least probability of rejecting the hypothesis if \( \mu \) is at the distance \( \Lambda \) or larger from the largest value in the hypothesis. Thus \( \alpha \) is what is called producer's risk in quality control problems and \( \beta \) the consumer's risk.

It has been proved [1], that no single-sample test satisfying (1) exists. It is possible, however, to devise a test employing two samples (the size of the first sample being fixed and the size of the second directed by the observations in the first) that insures the fulfillment of requirements (1) whatever be the value of \( \sigma \). The idea is due to Ch. STEIN [2]. In statistical quality control — as far as is known — Stein's method is not widely used, although with attribute sampling two-sample inspection plans (and indeed more complicated plans) are used occasionally. At the same time with the problem just stated the two-phase procedure is not merely a means for reducing the number of observations, but the only way to ensure the fulfillment of (1). The reason for which Stein's procedure has not so far been used more extensively seems to be that its application requires the computation of sample standard deviation which is not suitable for routine work. Therefore we propose a modification of Stein's procedure in which the usual estimate of standard deviation is replaced by mean range — an estimate that is common in quality control problems. In section 2 the modified procedure is described, in section 3 the theoretical background is explained and the construction of tables described. In section 4 the formulae for average sample size are developed and the average sample size tabulated for a particular case.
2. THE TEST PROCEDURE

The test procedure satisfying requirements (I) involves the following steps:

1. A random sample of an arbitrary size \( n_1 = kn \) is drawn (\( k \) and \( n \) are integers). Suitable choice of \( n_1 \) will be discussed in section 4. This sample is composed of \( k \) mutually independent subsamples of size \( n \). From this sample we compute: the mean, to be denoted by \( \bar{x}_1 \), the range of each subsample, \( R_1, R_2, \ldots, R_k \) and the mean range of all subsamples,

\[
\bar{R} = \frac{1}{k} \sum_{i=1}^{k} R_i.
\]

The sample is decomposed in subsamples for two reasons: first, for larger values of \( kn \) (say, \( kn \geq 12 \)) the mean range of \( k \) groups of \( n \) is — when \( k \) and \( n \) have been suitably chosen — a more efficient estimate of \( \sigma \) than the range of the total sample of \( kn \); second, it is easier to pick out the extreme observations in a smaller group than in the total sample.

2. Another sample is drawn the size of which, \( n_2 \), is given by the formula

\[
n_2 = \min \left\{ s : s > \frac{\bar{R}}{A^2Z^2} - n_1, \, s \text{ a positive integer} \right\}.
\]

In words, \( n_2 \) is the least positive integer greater than \( \frac{\bar{R}^2}{A^2Z^2} - n_1 \). Thus if \( \frac{R^2}{(A^2Z^2)} \leq n_1 \), we have \( n_2 = 1 \); if

\[
n_1 + s - 1 \leq \frac{\bar{R}^2}{A^2Z^2} < n_1 + s,
\]

we have \( n_2 = s \).

In formula (2) \( Z \) denotes a number depending on the chosen risks \( \alpha \) and \( \beta \),

\[
Z = \frac{1}{z_{0.05} + z_{0.10}}.
\]

The numbers \( z_{0.05} \) and \( z_{0.10} \) are for certain combinations of \( k \) and \( n \) given in tables 1 and 2, respectively. From the data in the second sample the mean \( \bar{x}_2 \) is computed.

3. The root \( a \) of the equation

\[
\frac{a^2}{n_1} + \frac{(1-a)^2}{n_2} = \frac{A^2Z^2}{\bar{R}^2}
\]

is determined.

The solution of (3) and the determination of \( n_2 \) from (2) are the only more complicated operations in the application of the procedure. However, if the test is to be used on repeated occasions with data regarding the same quality characteristic, suitable charts may be devised for these operations so that the computational labour is reduced to a minimum.
Table 1.
Values $z_{0.05}$ defined by the relation $P\left(\frac{\xi}{\sigma} > z_{0.05}\right) = 0.05$ ($\xi$ has a normal distribution with zero mean and unit variance, $w$ is distributed independently of $\xi$ as the mean range of $k$ mutually independent samples of size $n$ from a normal distribution with unit variance).

<table>
<thead>
<tr>
<th>$k$</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.869</td>
<td>0.765</td>
<td>0.698</td>
<td>0.648</td>
<td>0.616</td>
<td>0.588</td>
<td>0.566</td>
<td>0.544</td>
<td>0.510</td>
<td>0.464</td>
</tr>
<tr>
<td>2</td>
<td>0.780</td>
<td>0.703</td>
<td>0.654</td>
<td>0.612</td>
<td>0.583</td>
<td>0.560</td>
<td>0.541</td>
<td>0.525</td>
<td>0.490</td>
<td>0.452</td>
</tr>
<tr>
<td>3</td>
<td>0.754</td>
<td>0.684</td>
<td>0.635</td>
<td>0.600</td>
<td>0.573</td>
<td>0.551</td>
<td>0.533</td>
<td>0.518</td>
<td>0.484</td>
<td>0.448</td>
</tr>
<tr>
<td>4</td>
<td>0.742</td>
<td>0.675</td>
<td>0.629</td>
<td>0.594</td>
<td>0.568</td>
<td>0.547</td>
<td>0.529</td>
<td>0.515</td>
<td>0.482</td>
<td>0.446</td>
</tr>
<tr>
<td>5</td>
<td>0.735</td>
<td>0.660</td>
<td>0.624</td>
<td>0.591</td>
<td>0.565</td>
<td>0.544</td>
<td>0.527</td>
<td>0.513</td>
<td>0.480</td>
<td>0.445</td>
</tr>
<tr>
<td>6</td>
<td>0.730</td>
<td>0.666</td>
<td>0.622</td>
<td>0.589</td>
<td>0.563</td>
<td>0.543</td>
<td>0.526</td>
<td>0.511</td>
<td>0.479</td>
<td>0.444</td>
</tr>
<tr>
<td>7</td>
<td>0.726</td>
<td>0.663</td>
<td>0.620</td>
<td>0.587</td>
<td>0.562</td>
<td>0.542</td>
<td>0.525</td>
<td>0.510</td>
<td>0.478</td>
<td>0.444</td>
</tr>
<tr>
<td>8</td>
<td>0.724</td>
<td>0.662</td>
<td>0.618</td>
<td>0.586</td>
<td>0.561</td>
<td>0.541</td>
<td>0.524</td>
<td>0.510</td>
<td>0.478</td>
<td>0.443</td>
</tr>
<tr>
<td>9</td>
<td>0.722</td>
<td>0.660</td>
<td>0.617</td>
<td>0.585</td>
<td>0.560</td>
<td>0.540</td>
<td>0.523</td>
<td>0.509</td>
<td>0.477</td>
<td>0.443</td>
</tr>
<tr>
<td>10</td>
<td>0.721</td>
<td>0.659</td>
<td>0.616</td>
<td>0.584</td>
<td>0.559</td>
<td>0.539</td>
<td>0.523</td>
<td>0.509</td>
<td>0.477</td>
<td>0.443</td>
</tr>
<tr>
<td>12</td>
<td>0.718</td>
<td>0.657</td>
<td>0.615</td>
<td>0.583</td>
<td>0.558</td>
<td>0.539</td>
<td>0.522</td>
<td>0.508</td>
<td>0.476</td>
<td>0.442</td>
</tr>
<tr>
<td>15</td>
<td>0.716</td>
<td>0.656</td>
<td>0.613</td>
<td>0.582</td>
<td>0.558</td>
<td>0.538</td>
<td>0.521</td>
<td>0.507</td>
<td>0.476</td>
<td>0.442</td>
</tr>
<tr>
<td>20</td>
<td>0.714</td>
<td>0.654</td>
<td>0.612</td>
<td>0.581</td>
<td>0.557</td>
<td>0.537</td>
<td>0.521</td>
<td>0.507</td>
<td>0.475</td>
<td>0.441</td>
</tr>
<tr>
<td>30</td>
<td>0.712</td>
<td>0.652</td>
<td>0.611</td>
<td>0.580</td>
<td>0.556</td>
<td>0.536</td>
<td>0.520</td>
<td>0.506</td>
<td>0.475</td>
<td>0.441</td>
</tr>
<tr>
<td>60</td>
<td>0.709</td>
<td>0.651</td>
<td>0.610</td>
<td>0.579</td>
<td>0.555</td>
<td>0.535</td>
<td>0.519</td>
<td>0.505</td>
<td>0.474</td>
<td>0.441</td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.707</td>
<td>0.649</td>
<td>0.608</td>
<td>0.578</td>
<td>0.554</td>
<td>0.534</td>
<td>0.518</td>
<td>0.505</td>
<td>0.474</td>
<td>0.440</td>
</tr>
</tbody>
</table>

4. The hypothesis $\mu \leq \mu_0$ is accepted if the test statistic

$$T = \frac{a \bar{x}_1 + (1 - a) \bar{x}_2 - \mu_0}{\Delta \cdot Z}$$

(4)

214
does not reach the critical value $z_a$ from table 1 or 2 (according to the chosen risk $\alpha$) corresponding to the pair $k$ and $n$ that has been employed in subdividing the size of the first sample, and rejected otherwise.

By a modification of the inequalities we obtain a test satisfying the conditions

$$P(\mu_0) = \alpha,$$
$$P(\mu_0 - \Lambda) = \beta,$$

i.e. a test of $\mu \geq \mu_0$ directed against the alternatives

$$\mu < \mu_0.$$

3. THEORETICAL BACKGROUND

Let us examine the distribution of the random variable

$$\zeta = \frac{a x_1 + (1 - a) x_2 - \mu}{\Lambda Z},$$

where $x_1$ and $x_2$ are the means of two random samples from a normal population with mean $\mu$ and standard deviation $\sigma$ taken in accordance with the principles enumerated in section 2, $a$ is a function of the first sample given by equation (3), $\Lambda$ and $Z$ have the same meaning as in section 2. First we consider the conditional distribution of $\zeta$, given $\bar{X}$. This distribution is normal (since it is a linear combination of means of two normal samples) and has the expectation

$$\frac{1}{\Lambda Z} E \{a x_1 + (1 - a) x_2 - \mu\} = 0$$

and variance

$$\frac{\sigma^2}{\Lambda^2 Z^2} \left\{ \frac{a^2}{n_1} + \frac{(1 - a)^2}{n_2} \right\} = \frac{\sigma^2}{\bar{R}^2}$$

(the last equality follows from (3)).

Thus the conditional distribution function of $\zeta$ given $\bar{R}$ is

$$P_R(z) = P(\zeta \leq z \mid \bar{R}) = \Phi \left( \frac{z \bar{R}}{\sigma} \right),$$

where $\Phi$ denotes the distribution function of the normal distribution with zero mean and unit variance. And the unconditional distribution function of the random variable $\zeta$ is

$$P(z) = P(\zeta \leq z) = \int_0^{\infty} \Phi(zv) p_{k,n}(v) \, dv,$$  \hspace{1cm} (5)

where $p_{k,n}(v)$ denotes the probability density of the distribution of mean range of $k$ mutually independent samples of size $n$ from a normal population with unit variance. It may be easily verified that (5) is the distribution function
of the ratio of two random variables, say $X/Y$, where $X$ has a normal distribution with zero mean and unit variance and $Y$ is distributed as mean range of $k$ independent samples of size $n$ from a normal population with unit variance. Also, this is the distribution of the test statistic $T$ under the hypothesis $\mu = \mu_0$.

The numbers $z_q$ in tables 1 and 2 are the solutions of the equation

$$1 - P(z_q) = q$$

for $q = 0.05$ and $q = 0.10$, respectively.

It is seen that the procedure described in section 2 satisfies the requirements (1) of section 1 since

$$P(\mu_0) = P(T \geq z_a \mid \mu = \mu_0) = 1 - P(z_a) = \alpha$$

and

$$P(\mu_0 + A) = P(T \geq z_a \mid \mu = \mu_0 + A) = 1 - P\left(z_a - \frac{1}{Z}\right) = 1 - P(-z_\beta) = 1 - \beta.$$

The last equality follows from the symmetry of the distribution of $\zeta$.

The values in tables 1 and 2 have been computed by means of an approximation suggested by P. B. Patnaik [3]. Numerical comparisons which have been made so far show that the approximation is very close. We shall describe it briefly, since we shall have occasion to use it in section 4 in the calculation of the average number of necessary observations.

Patnaik's approximation consists in the replacement of the distribution of mean range by the distribution of the random variable $c\sqrt{\chi^2/v}$, where $\chi^2$ has a $\chi^2$ distribution with $r$ “degrees of freedom” and $c$ is a positive constant. The constants $c$ and $r$ are determined so that the first two moments of $c\sqrt{\chi^2/v}$ coincide with the first two moments of mean range. Thus the “number of degrees of freedom” may assume non-integer values, too. The numbers $c$ and $r$ depend, of course, on the number of subsamples, $k$, and on their size, $n$. A table of $c$ and $r$ values for some combinations of $k$ and $n$ is given in [4]. For the purposes of the present paper a more detailed table had to be prepared.

The critical values $z_q$ were then computed as follows. Let $\xi$ have a normal distribution with zero mean and unit variance, further let $w$ be distributed as mean range of $k$ independent samples of size $n$ from a normal population with unit variance; assume that $w$ is independent of $\xi$. Since $w$ is distributed approximately as $c\sqrt{\chi^2/v}$, the ratio $\xi/w$ is distributed approximately as the ratio $\xi/[c\sqrt{\chi^2/v}]$, i.e.

$$P\left\{\frac{\xi}{w} \leq z\right\} \approx P\left\{\frac{\xi}{c\sqrt{\chi^2/v}} \leq z\right\} = G(cz),$$

(7)
where \( G \) denotes the distribution function of Student's distribution with \( v \) "degrees of freedom". The values \( z_q \) defined by (6) are thus approximately given by

\[
    z_q = \frac{t_{1-q}}{c} ,
\]

where \( t_{1-q} \) is the 100(1 - \( q \)) percent fractile of Student's distribution with \( v \) "degrees of freedom".

For some particular values of \( q \) the critical values \( z_q \) may be computed from the values tabulated by Lord in [5]. Lord's tables are constructed for a two-sided test of Student's hypothesis when an unbiased estimate of \( \sigma \) based on mean range is used. Accordingly they give, for selected values of \( q \) and for a number of combinations of \( k \) and \( n \) (i.e. the number of subsamples and their size), the values \( u_q \) defined by the relation

\[
    P \left\{ \frac{|\xi|}{w/d_n} > u_q \right\} = q ,
\]

where \( \xi \) is a random variable with a normal distribution with zero mean and unit standard deviation and \( w \) is a random variable that is distributed as mean range of \( k \) mutually independent samples of size \( n \) from a normal population with unit variance; \( d_n \) is the expectation of the range of a sample of \( n \) elements from a unit normal population, tabled e.g. in [4] or [6]. The critical values \( z_q \) are connected with \( u_q \) through the relation

\[
    z_q = \frac{u_{2q}}{d_n} .
\]

4. THE AVERAGE SAMPLE SIZE

In the procedure described in section 2 only the first sample has a fixed size \( n_1 \); the size of the second sample, \( n_2 \), depends on the mean range of the first sample. The total number of observations necessary to reach the decision, \( N = n_1 + n_2 \), is thus, in fact, a random variable. Its mean value is

\[
    \mathbb{E}\{N\} = (n_1 + 1) \mathbb{P}\{\bar{R} < A \bar{Z} | \bar{R}_1 + 1\} + \sum_{k=2}^{\infty} k \mathbb{P}\left\{ k - 1 \leq \frac{\bar{R}^2}{A^2Z^2} < k \right\} =
\]

\[
= (n_1 + 1) \mathbb{P}\left\{ \frac{\bar{R}^2}{\sigma^2} < \frac{A^2Z^2(n_1 + 1)}{\sigma^2} \right\} + \sum_{k=2}^{\infty} k \mathbb{P}\left\{ \frac{A^2Z^2(k - 1)}{\sigma^2} \leq \frac{\bar{R}^2}{\sigma^2} < \frac{A^2Z^2k}{\sigma^2} \right\} .
\]

The meaning of the symbols \( \bar{R}, A, Z \) and \( \sigma^2 \) is the same here as in sections 1 and 2. Denote by \( \delta \) the "distance between the hypothesis and the alternative", expressed in terms of the population standard deviation, \( \delta = A/\sigma \) and apply Patnaik's approximation described in section 3 to the expression for \( \mathbb{E}\{N\} \).
We have

\[ E\{N\} \approx (n_1 + 1) \left( 1 + \frac{\delta^2 Z_2^2(n_1 + 1)}{c^2} \right) \]

\[ + \sum_{k=n_1+2}^{\infty} k \int \frac{\delta^2 Z_2^2(k-1)(k-1)}{c^2} \left\{ \frac{x^2}{c^2} \leq \chi^2_x \right\} \]

where \( \chi^2_x \), \( v \) and \( c \) have the same meaning as in section 3. If we denote by \( F_v \) the distribution function of the \( \chi^2 \)-distribution with \( v \) "degrees of freedom", we may write the expectation \( E\{N\} \) as

\[ E\{N\} \approx (n_1 + 1) F_v(A_{n_1+1}) + \sum_{k=n_1+2}^{\infty} k \int \frac{A_k}{A_{k-1}} \ dF_v(x) , \]

where \( A_k \) stands for \( (\delta^2 Z^2 k^2)/c^2 \). In the second term on the right-hand side we now replace the summands by \( \int \frac{c^2}{\delta^2 Z^2 v} \frac{A_k}{A_{k-1}} x \ dF_v(x) \), thus introducing a further approximation. We obtain

\[ E\{N\} \approx (n_1 + 1) F_v(A_{n_1+1}) + \frac{c^2}{\delta^2 Z^2 v} \int \frac{A_k}{A_{k-1}} x \ dF_v(x) . \]

Integrating by parts we get finally

\[ E\{N\} \approx (n_1 + 1) F_v(A_{n_1+1}) + \frac{c^2}{\delta^2 Z^2 v} \left[ 1 - F_{v+2}(A_{n_1+1}) \right] . \] (10)

If the first sample is large, the number \( v \) will usually be large, too, and the expression (10) may be replaced by

\[ E\{N\} \approx (n_1 + 1) \left( 1 - \frac{\delta^2 Z_2^2}{c^2} (2^n - 1) \right) + \frac{c^2}{\delta^2 Z^2} \left[ 1 - \Phi(\sqrt{2A_{n_1+1}} - \sqrt{2v + 3}) \right] , \]

where \( \Phi \) denotes the distribution function of the normal distribution with zero mean and unit variance.

Thus the mean number of observations depends on the population standard deviation \( \sigma \), on the size of the first sample and on the way of its subdivision into subsamples, i.e. on \( k \) and \( n \). It is natural to choose the size \( n_1 \) and its subdivision so as to make the mean number of observations as small as possible. To this end it is necessary to have some idea of the value of the population standard deviation, to fix the size of subsamples into which the first sample will be divided, to compute \( E\{N\} \) using (10) or (11) for several values of \( n_1 \) and to minimize it by trial and error. Numerical comparison shows that the dependence of the expected number of observations upon the method of the subdivision of the first sample is not too strong. Table 3 gives values \( E\{N\} \) corresponding to various choices of \( k \) and \( n \) for \( n_1 = kn = 100 \) and to various
values of $\delta$, when the risks are $\alpha = 0.05$ and $\beta = 0.10$. For $\delta = 0.4$ the expected number of observations is approximately 101, so that for $\delta$ as high as 0.4 the second sample will usually be of size 1, it will consist of a single observation. This means that for $\delta \geq 0.4$ the first sample of size 100 is unnecessarily large (in this connection see Table 4 and the corresponding comments). Further it is seen from Table 3 that the expected number of observations is at its minimum value when the subsamples have sizes somewhere between 5 and 15. This is in agreement with the well known fact that the mean range as an estimate of the standard deviation is most efficient when the size of subsamples is about 8.

<table>
<thead>
<tr>
<th>Number of subsamples $k$</th>
<th>Size of subsample $n$</th>
<th>Difference between hypothesis and alternative in terms of population $x$, d, $\sigma$: $\delta = 0.02$</th>
<th>$\delta = 0.05$</th>
<th>$\delta = 0.10$</th>
<th>$\delta = 0.20$</th>
<th>$\delta = 0.30$</th>
<th>$\delta = 0.40$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>20</td>
<td>21954</td>
<td>3513</td>
<td>878.2</td>
<td>219.5</td>
<td>106.2</td>
<td>101.0</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>21863</td>
<td>3498</td>
<td>874.5</td>
<td>218.6</td>
<td>102.5</td>
<td>101.0</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>21897</td>
<td>3503</td>
<td>875.9</td>
<td>219.0</td>
<td>102.7</td>
<td>101.0</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
<td>22234</td>
<td>3557</td>
<td>889.3</td>
<td>222.3</td>
<td>108.2</td>
<td>101.0</td>
</tr>
<tr>
<td>Size of single sample for $\sigma$ known</td>
<td>21418</td>
<td>3427</td>
<td>857</td>
<td>214</td>
<td>95</td>
<td>54</td>
<td></td>
</tr>
</tbody>
</table>

Thus it is sufficient for the determination of the optimal $n_1$ to compute $E\{N\}$ for $n_1$ growing by steps of 10, beginning with $n_1$ equal to about one half the size of sample necessary for the attainment of the same power with the single-sample test with known standard deviation, and to use subsamples of size 10. In Table 3 the sizes for the single-sample test with the same power (the same risks) when the standard deviation is known are shown in the last row.

For illustration which $n_1$ is approximately optimum we give another table, Table 4, which gives the values $E\{N\}$ corresponding to various sizes of the first sample, $n_1$, and various values of $\delta$, when $\alpha = 0.05$ and $\beta = 0.10$ as before. It is assumed that the first sample is divided into subsamples of size 10. Tables 4 and 3 have been computed by means of formulas (10) and (11).

The last two rows of Table 4 give, respectively, the size of the sample necessary for the single-sample test having the same power $1 - \beta$ against the alternatives specified by $\delta$ for the case when $\sigma$ is known, and for the case when $\sigma$ is estimated by means of the common mean-square estimate (these values are quoted from [7]). It is seen that the minimal expected number of observations does not exceed much the size necessary for the single-sample test with
Table 4.
Expected number of observations for the two-sample procedure with various sizes of the first sample.

<table>
<thead>
<tr>
<th>Size of first sample</th>
<th>Difference between hypothesis and alternative in terms of population s. d. σ:</th>
</tr>
</thead>
<tbody>
<tr>
<td>n₁</td>
<td>δ 0-25 0-30 0-35 0-40 0-45 0-50 0-60</td>
</tr>
<tr>
<td>10</td>
<td>171-2 118-9 87-4 66-9 52-9 42-9 29-9</td>
</tr>
<tr>
<td>20</td>
<td>152-9 106-2 78-0 59-8 47-3 38-5 27-9</td>
</tr>
<tr>
<td>30</td>
<td>147-3 102-3 75-2 57-7 46-1 38-6 32-2</td>
</tr>
<tr>
<td>40</td>
<td>144-7 100-4 73-8 57-1 47-4 43-0 41-1</td>
</tr>
<tr>
<td>50</td>
<td>143-0 99-3 73-3 58-7 52-7 51-2 51-0</td>
</tr>
<tr>
<td>60</td>
<td>142-1 98-7 74-1 63-6 61-3 61-0 61-0</td>
</tr>
<tr>
<td>70</td>
<td>141-6 98-6 77-2 71-6 71-0 71-0 71-0</td>
</tr>
<tr>
<td>80</td>
<td>140-8 99-1 83-1 81-1 81-0 81-0 81-0</td>
</tr>
<tr>
<td>90</td>
<td>140-4 101-2 91-3 91-0 91-0 91-0 91-0</td>
</tr>
<tr>
<td>100</td>
<td>140-2 105-7 101-1 101-0 101-0 101-0 101-0</td>
</tr>
<tr>
<td>120</td>
<td>141-2 121-5 121-0 121-0 121-0 121-0 121-0</td>
</tr>
<tr>
<td>150</td>
<td>154-2 151-0 151-0 151-0 151-0 151-0 151-0</td>
</tr>
</tbody>
</table>

Size of single sample for σ known:  

<table>
<thead>
<tr>
<th>Size of single sample for σ known:</th>
</tr>
</thead>
<tbody>
<tr>
<td>137 93 70 34 42 34 24</td>
</tr>
</tbody>
</table>

Size of single sample, σ estimated by mean-square estimate:  

<table>
<thead>
<tr>
<th>Size of single sample, σ estimated by mean-square estimate:</th>
</tr>
</thead>
<tbody>
<tr>
<td>139 97 72 55 44 36 26</td>
</tr>
</tbody>
</table>

Known σ. Further it is at first sight surprising that no reduction in the number of observations results from the decomposition of the test into two phases in the case of unknown standard deviation. It must be remembered, however, that the described two-sample procedure is devised to satisfy one condition more, viz. the given power against A irrespective of σ. If this condition were replaced and the requirement on the power stated only in terms of δ (that is in terms of σ), then, it is believed, it would be possible to find a more economical procedure. This, however, constitutes another problem.

References

TEST STUDENTOVY HYPOTHESY DVOJÍM VÝBĚREM PŘI UŽITÍ PRŮMĚRNÉHO ROZPĚTÍ

JOSEF MACHEK

(Došlo dne 9. září 1958.)

V článku je popsána modifikace Steinova dvojřádového testu hypotézy o průměru normálního rozdělení, při které se na místo obvyklého odhadu směrodatné odchylky užívá výběrového rozpětí, resp. průměrného rozpětí několika výběrů.

Podstata úlohy je tato: Je dán základní soubor, jenž má normální rozdělení s neznámou střední hodnotou \( \mu \) a s neznámou směrodatnou odchylkou \( \sigma \). Jest ověřiti hypotézu \( H_0: \mu = \mu_0 \) proti alternativě \( H_1: \mu > \mu_0 \). Test má splňovat požadavky (1), kde \( P(\mu) \) značí pravděpodobnost, že testovaná hypotéza \( H_0 \) bude zamítnuta, když skutečná hodnota průměru je \( \mu \); \( A \) je dané číslo (předpokládáme \( A > 0 \)) a \( \alpha \) a \( \beta \) jsou zvolená risika chybných rozhodnutí. Požadavky (1) mají být splněny nezávisle na hodnotě směrodatné odchylky \( \sigma \).

Je dokázáno, že neexistuje test, založený na jediném výběru, který by měl požadované vlastnosti (1). Lze však sestrojit test, založený na dvou výběrech, z nichž druhý má rozsah závislý na pozorováních z prvního výběru, který splňuje (1) při jakémkoliv \( \sigma \) (2).

Při užití rozpětí na místě obvyklého odhadu směrodatné odchylky zahrnuje test tyto kroky:

1. Vezme se výběr libovolného rozsahu \( n_1 = kn \) (kde \( k \) a \( n \) jsou celá čísla). Výběr je složen z \( k \) navzájem nezávislých dílčích výběrů rozsahu \( n \). Vypočte se aritmetický průměr celého prvního výběru, rozpětí každého z dílčích výběrů a průměrné rozpětí \( R \) všech dílčích výběrů.

2. V závislosti na zvolených hodnotách \( \alpha \) a \( \beta \) se určí číslo \( Z = \frac{1}{z_\alpha + z_\beta} \), kde \( z_\alpha \) jsou pro \( q = 0,05 \) a \( q = 0,10 \) uvedena v tabulkách 1 a 2. Provede se...
druhý výběr, jehož rozsah \( n_2 \), je dán výrazem (2), a vypočte se jeho aritmetický průměr \( \bar{x}_2 \).

3. Určí se konstanta \( a \) řešením rovnice (3).

4. Vypočte se hodnota testové charakteristiky (4) a testovaná hypotéza se zamítne, jestliže \( T \geq z_a \). Kritické hodnoty testové charakteristiky jsou uvedeny v tabulkách 1 a 2.

V § 3 je odvozeno rozdělení testové charakteristiky; charakteristika \( T \) má rozdělení jako podíl dvou navzájem nezávislých náhodných veličin, \( X/Y \), kde \( X \) má rozdělení normální s parametry \( \theta \) a \( 1 \) a \( Y \) má rozdělení jako průměrné rozpětí \( Ic \) navzájem nezávislých výběrů rozsahu \( n \) ze základního souboru s normálním rozdělením s jednotkovým rozpětlem.

Při popsaném postupu je rozsah druhého výběru, a tedy i celkový počet pozorování, \( N \), nutný k dosažení rozhodnutí, náhodnou veličinou. Její přibližná střední hodnota je dána výrazy (10) a (11). Závisí na směrodatné odchylce \( \sigma \) základního souboru. Pomocí formulí (10) a (11) byla vypočtena tabulka 4, udávající střední hodnoty počtu pozorování při různých volbách \( n_1 \) a při různých hodnotách \( \delta = \Delta/\sigma \). Formulí (10) a (11) lze užít při volbě optimálního \( n_1 \), tj. takového, při kterém střední hodnota celkového rozsahu výběru je minimální.

Резюме

КРИТЕРИЙ ДЛЯ ПРОВЕРКИ ГИПОТЕЗЫ СТЮДЕНТА
ДВОЙНОЙ ВЫБОРКОЙ

Иосиф Макек (Josef Machek)

(Поступило в редакцию 9/IX 1958 г.)

В статье рассматривается следующая задача: Распределение генеральной совокупности нормально с неизвестными параметрами \( \mu \) (среднее значение) и \( \sigma \) (стандартное отклонение). Требуется проверить гипотезу \( \mu = \mu_0 \) против альтернативной гипотезы \( \mu > \mu_0 \). Обозначим через \( P(\mu) \) вероятность непринятия проверяемой гипотезы, если истинное значение среднего равно \( \mu \). Критерий для проверки должен удовлетворять требованиям (1), где \( \Delta \) — заданное число и \( \alpha \) и \( \beta \) — заранее выбранные вероятности ошибочных решений, независимо от неизвестного значения параметра \( \sigma \).

Известно, что при помощи одной выборки эта задача не может быть решена. Для ее решения необходимо впользоваться двумя выборками, причем объем второй из них является функцией наблюдений, полученных в первой выборке. Этот метод принадлежит X. Штейну [2]. В настоящей
статье метод Штейна видоизменен так, что в качестве оценки неизвестного параметра использован размах выборки или средний размах нескольких выборок.

Проверка включает следующие шаги:

1. Из совокупности производят выборку произвольного объема \( n_1 = kn \), где \( k \) и \( n \) — целые числа. Эту выборку разобьют на \( k \) частичных выборок объема \( n \) и затем подсчитывают среднее всей выборки \( \bar{x}_1 \), размах каждой из частичных выборок и средний размах всех \( k \) частичных выборок \( R \).

2. Возьмут вторую выборку, объем которой \( n_2 \) задан формулой (2), где \( Z = (z_0 + z_0) - 1 \); значения \( z_{0.05} \) и \( z_{0.10} \) для различных сочетаний \( k \) и \( n \) приведены в таблицах 1 и 2. По данным второй выборки подсчитывают среднее \( \bar{x}_2 \).

3. Подсчитывают значение критерия \( T \), заданного формулой (4), где \( \alpha \) есть решение уравнения (3). Гипотеза \( H_0 \) отвергнута, если \( T \geq z_\alpha \), где \( \alpha \) — заранее фиксированная уровень значимости.

В разделе 3 приведено распределение критерия \( T \) при гипотезе и при альтернативах.

При только что описанном критерии число наблюдений, необходимых для решения, является случайной величиной. Ее математическое ожидание приблизительно задано формулами (10) или (11), где \( F_v \) обозначает функцию распределения величины \( \chi^2 \) с \( v \) степенями свободы, \( c \) — постоянная, зависящая от \( k \) и \( n \) (\( c \) и \( v \) табулированы, например, в [4]), \( \delta = \Delta/\sigma \), \( \Phi \) — функция нормального распределения с параметрами 0 и 1. По формулам (10) и (11) была составлена таблица 4, в которой приведены математические ожидания числа наблюдений для некоторых значений \( \delta \) и \( n_1 \). Формулы (10) и (11) могут быть использованы для оптимального выбора объема первой выборки, дающего минимальное среднее число наблюдений.