# RESEARCH ARTICLES

*J. Sci. Soc. Thailand*, **5** (1979) 73-87

## THE WATER STORAGE PROBLEM WITH INDEPENDENT NORMAL INFLOWS

H.N. PHIEN and A. ARBHABHIRAMA

*Asian Institute of Technology, P.O. Box 2754, Bangkok, Thailand*

P. SUTABUTR

*National Energy Administration of Thailand, Bangkok, Thailand*

*(Received 1 March 1979)*

**Summary**

*Assuming that the annual inflows to the reservoir under consideration are distributed as independent normal variables and the annual outflows are equal to the sample mean of the inflows, the distributions of the water content in various years, the maximum amounts of water in surplus and in deficit, and the storage capacity were investigated under the condition that the reservoir allows neither spillage nor emptiness during its lifetime. The water contents in various years were found to be distributed as normal variables; the maximum amount of water in surplus or in deficit has a probability mass at zero and a probability density function elsewhere, which was fitted by the Type I Pearson curve; and the storage capacity was fitted by the Type III curve. All the fittings were assessed by using the Chi-square test of goodness of fit.*

## Introduction

The planning and design of a reservoir always involves the determination of its storage capacity, which depends upon the conditions imposed. Suppose that the annual flows to the reservoir in different years, denoted by $X_k$, come from the same population distribution. Suppose furthermore that during a lifetime of n years, the annual outflows are constants equal to the sample mean $\overline{X}_n$:

$$\overline{X}_n = \sum_{k=1}^{n} X_k/n \qquad (1)$$

The problem is to determine the reservoir size or the reservoir storage capacity in order that during the lifetime the reservoir never spills and never runs dry.

Let $S_i^*$ denote the storage of the reservoir or the water content in the ith year with reference to the initial level, for $i = 0, 1, \ldots, n$, where $S_0^* = 0$, then by means of the continuity equation

$$S_i^* - S_{i-1}^* = X_i - \overline{X}_n$$

it follows that

$$S_i^* = \sum_{k=1}^{i} (X_i - \overline{X}_n)$$

Let $S_i = \sum_{k=1}^{i} X_k$, the partial sum of $X_k$; it can be seen that

$$S_i^* = S_i - \frac{i}{n} S_n, \; i = 1, 2, \ldots, n \tag{2}$$

In this form, $S_i^*$ is called the adjusted partial sum.

Define

$$M_n^* = \max (0, S_1^*, \ldots, S_n^*)$$
$$m_n^* = \min (0, S_1^*, \ldots, S_n^*) \tag{3}$$

and

$$R_n^* = M_n^* - m_n^*,$$

$M_n^*$ is called the adjusted surplus, $m_n^*$ the adjusted deficit, and $R_n^*$ the adjusted range. To satisfy the stated conditions, the storage capacity of the reservoir should be optimally equal to the adjusted range $R_n^*$. The random variables $M_n^*$ and $m_n^*$ are then the maximum amounts of water in surplus and in deficit over n years, respectively. The random variables $S_i^*$, $M_n^*$, $m_n^*$ and $R_n^*$, illustrated in Fig. 1 are therefore of important interest in the field of water resources development.

Hurst[1] collected a large amount of statistical material relating to water levels and other phenomena, and came to the conclusion that the observed adjusted range has a mean proportional to $n^c$, where c varies from 0.69 to 0.80, with a mean of 0.729 and a standard deviation of 0.092. This finding, commonly referred to as the Hurst phenomenon, has been a stimulant for a great deal of research. Recently, it has been considered to be related to the rescaled adjusted range[2].

Feller[3] derived the asymptotic distribution of $R_n^*$ and its asymptotic mean and variance. For independent variables of zero mean and unit variance, the results are

$$E(R_n^*) \doteq (\pi n/2)^{1/2} \tag{4}$$

$$Var(R_n^*) \doteq \pi(\pi/6-1/2)n \tag{5}$$

Solari and Anis[4] introduced the random variable $U_n^*$ called the maximum of adjusted partial sums which is defined by

$$U_n = \max (S_1^*, \ldots, S_n^*)$$

and derived the exact formulas for its first two moments as follows:

$$E(U_n^*) = \frac{1}{2} \left(\frac{n}{2\pi}\right)^{1/2} \sum_{i=1}^{n-1} i^{-1/2} (n-i)^{-1/2} \tag{6}$$

$$E(U_n^{*2}) = \frac{1}{6} \left\{ \frac{n^2-1}{n} + \frac{\sqrt{n}}{2\pi} \sum_{i=2}^{n-1} \sum_{j=1}^{i-1} i(2i-n) \left[(n-i)j^3 (i-j)^3\right]^{-1/2} \right\} \tag{7}$$

for $n \geq 2$

Salas-La Cruz[5] obtained the expected value of the adjusted range for the case of independent standard normal variables. The result was:

$$E(R_n^*) = \left(\frac{n}{2\pi}\right)^{1/2} \sum_{i=1}^{n-1} \frac{2(n-i)^{1/2}}{ni^{1/2}} \tag{8}$$

Later, Boes and Salas-La Cruz[6] and Salas-La Cruz and Boes[7] derived the expected value of $R_n^*$ for the case where the outflow is only a fraction of the sample mean $\overline{X}_n$. They also investigated the asymptotic behaviour of $E(R_n^*)$ in connection with the Hurst phenomenon.

Sutabutr[8], based on computed values of the Pearson criterion from simulated samples, suggested the use of the Type III curve of the Pearson system to fit the distribution of $R_n^*$.

As seen before, the random variables adjusted surplus, adjusted deficit, and adjusted range are very important in the storage problem. An attempt is made in this study to obtain their distributions when $X_k$ are distribution as independent normal variables.

## Expected Values of $M_n^*$, $m_n^*$ and $R_n^*$

It should be noted that for $n = 1$, $\overline{X}_n = X_1$, hence $S_1^* = 0$, and consequently $M_1^* = m_1^* = R_1^* = 0$. This case is excluded from the following analysis. For $n \geq 2$, the sequence $Y_1, Y_2, \ldots, Y_n$ where $Y_k = X_k - \overline{X}_n$ is exchangeable[6]; thus the following equation[9] holds

$$E(M_n^*) = \sum_{i=1}^{n} E(V_i)/i \tag{9}$$

where

$$V_i = \max (0, S_i^*) = |S_i^*|/2 + S_i^*/2 \tag{10}$$

Observe that since

$$\begin{aligned} m_n^* &= \min (0, S_1^*, \ldots, S_n^*) \\ &= -\max (0, -S_1^*, \ldots, -S_n^*) \end{aligned}$$

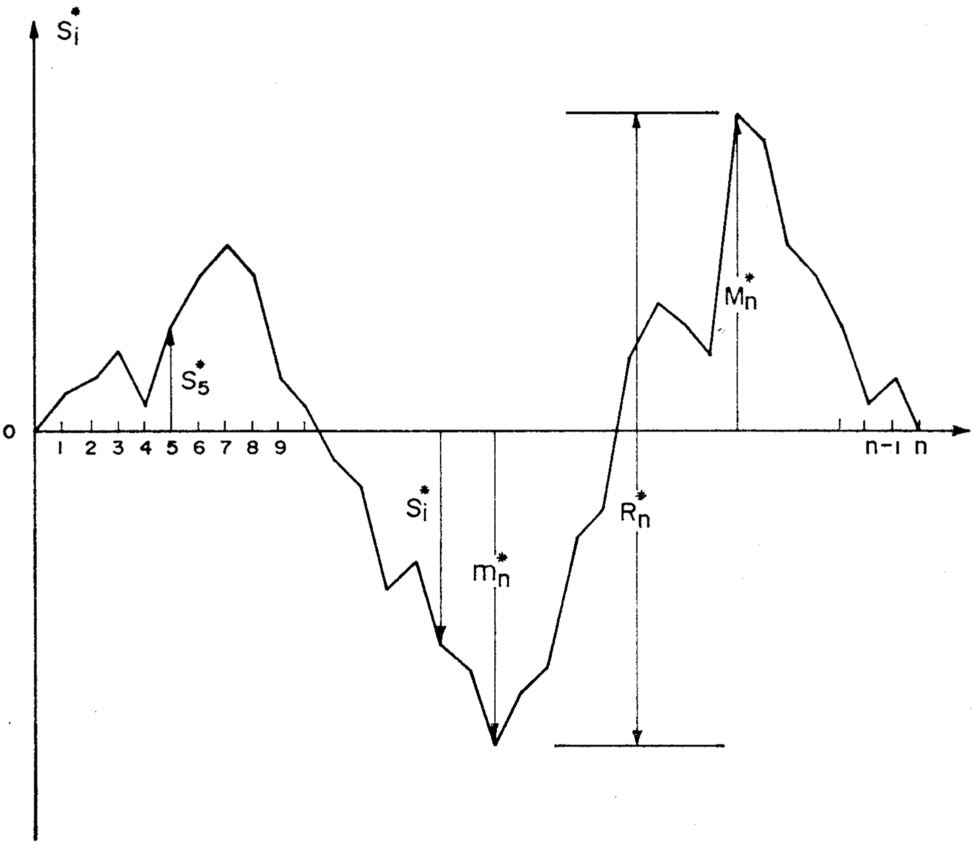


**Fig. 1.** Definition of the adjusted partial sum $S_i^*$, adjusted surplus $M_n^*$, adjusted deficit $m_n^*$. and adjusted range $R_n^*$.

it follows that

$$E(m_n^*) = -\sum_{i=1}^{n} E(W_i)\ /i \tag{11}$$

where

$$W_i = \max\ (0, -S_i^*) = |S_i^*|\ /2 - S_i^*\ /2 \tag{12}$$

From the definition of $S_i^*$,

$$E(S_i^*) = 0$$

thus,

$$E(V_i) = E(W_i) = E(|S_i^*|)\ /2 \tag{13}$$

and from (9) and (11),

$$E(M_n^*) = -E(m_n^*) \tag{14}$$

The definition of $R_n^*$ leads to:

$$E(R_n^*) = E(M_n^*) - E(m_n^*)$$

or

$$E(R_n^*) = 2E(M_n^*) \tag{15}$$

Equations (14) and (15) are abtained using the fact that $E(S_i^*) = 0$, which in turn results from the condition $E(Y_k) = 0$, $k = 1, ..., n$, where $Y_k = X_k - \overline{X}_n$ is called the net input in the kth year.

These equations state that the expected values of the adjusted surplus and adjusted deficit have the same magnitude, and the expected value of the adjusted range is double of the adjusted surplus, provided that the expected value of the net inputs in different years are identically equal to zero.

For the case of independent normal inflows with mean $\mu$ and standard deviation $\sigma$.

$$E(S_i) = \sum_{k=1}^{i} E(X_k) = i\mu,$$

and

$$Var(S_i) = \sum_{k=1}^{i} Var(X_k) = i\sigma^2$$

The expected value of $S_i^*$ is zero, and the variance can be written, using (2), as

$$Var(S_i^*) = Var(S_i) + (i/n)^2\ Var(S_n) - 2i\ Cov(S_i, S_n)\ /n$$

Since

$$Cov(S_i, S_n) = i\sigma^2,$$

$$Var(S_i^*) = i(n-i)\sigma^2\ /n. \tag{16}$$

It should be noted that $\overline{X}_n$ is a normal variable, so is $Y_k$, and therefore $S_i^*$ is a normal variable. The knowledge of the expected value and variance of $S_i^*$ thus completely determines the distribution of $S_i^*$.

In summary, the storage of the reservoir in the ith year, $i \leq n$, is distributed as a normal variable with zero mean, and variance given by (16). The expected value of $|S_i^*|$ can be easily shown to be

$$E(|S_i^*|) = (2/\pi)^{1/2}\ [i(n-i)/n]^{1/2}\ \sigma$$

Substituting this expresssion into (13) yields

$$E(V_i) = E(W_i) = (2\pi)^{-1/2} [i(n-i)/n]^{1/2}\sigma \quad (17)$$

The expected values of $M_n^*$ and $m_n^*$ are obtained using (9) and (11), respectively and the expression of $E(V_i)$ and $E(W_i)$ in (17):

$$E(M_n^*) = -E(m_n^*)$$

$$= (2\pi)^{-1/2}\sigma \sum_{i=1}^{n} \left(\frac{n-i}{ni}\right)^{1/2} \quad (18)$$

and the expected value of $R_n^*$ is given by (15)

$$E(R_n^*) = (2/\pi)^{1/2} \sigma \sum_{i=1}^{n} \left(\frac{n-i}{ni}\right)^{1/2} \quad (19)$$

**Second Moments of $M_n^*$, $m_n^*$ and $R_n^*$**

For $n = 2$,

$M_2^* = \max(0, S_1^*, S_2^*) = V_1$

because $S_2^* = 0$. Thus

$$E(M_2^{*2}) = E(V_1^2) = E(S^{*2})/2 = \sigma^2/4 \quad (20)$$

Similarly, $m_2^* = -W_1$, and

$$E(m_2^{*2}) = E(W_1^2) = E(S_1^{*2})/2 = \sigma^2/4 \quad (21)$$

For $n \geq 3$, the second moments of $M_n^*$ and $m_n^*$ can be obtained from the formula of Solari and Anis[4], by noting that the random variable

$U_n^* = \max(S_1^*, S_2^*, \ldots, S_{n-1}^*, S_n^*)$

corresponding to the condition $S_n^* = 0$ becomes

$$U_n^* = \max(S_1^*, S_2^*, \ldots, S_{n-1}^*, 0) = M_n^* \quad (22)$$

The second moment of $M_n^*$ is then derived from (7)

$$E(M_n^{*2}) = \frac{\sigma^2}{6}\left\{\frac{n^2-1}{n} + \frac{\sqrt{n}}{2\pi}\sum_{i=1}^{n-1}\sum_{j=1}^{i-1} i(2i-n)\,[(n-i)j^3\,(i-j)^3]^{-1/2}\right\} \quad (23)$$

for $n \geq 3$.

Equation (23) obviously can apply to the second moment of $m_n^*$ for normal variables. Based on the identification between $U_n^*$ and $M_n^*$, one can show that the expected value of $M_n^*$ derived in (18) is equivalent to the expected value of $U_n^*$ obtained by Solari and Anis[4] in (6) for the case $\sigma = 1.0$.

The exact formula for the second moment of $R_n^*$ can be obtained for $n = 2$. It is clear that $R_2^* = S_1^*$, thus

$$E(R_2^{*2}) = \sigma^2/2 \quad (24)$$

For $n \geq 3$, the Monte Carlo method was used to compute the variance of $R_n^*$ from simulated samples. It was found that for a size of 5,000, the values of $E(M_n^{*2})$

**TABLE I: VARIANCE OF $R_n^*$ FOR THE CASE OF INDEPENDENT NORMAL VARIABLES ($\sigma = 1.0$)**

| n | From approximate formula | Simulated values |
|---|---|---|
| 4 | 0.3589 | 0.3672 |
| 6 | 0.5071 | 0.5063 |
| 8 | 0.6553 | 0.6491 |
| 10 | 0.8035 | 0.8145 |
| 15 | 1.1740 | 1.1738 |
| 20 | 1.5445 | 1.5530 |
| 25 | 1.9150 | 1.9097 |
| 30 | 2.2855 | 2.2863 |
| 35 | 2.5660 | 2.6596 |
| 40 | 3.0265 | 3.0198 |
| 45 | 3.3970 | 3.3985 |
| 50 | 3.7675 | 3.7662 |

computed in this method were close to those computed from the exact formula of (23), thus for each n, a sample of size equal to 5,000 was used. From the computed values, the variance of $R_n^*$ was found to be approximated by the following equation:

$$\mathrm{Var}(R_n^*) = (0.0741n + 0.0625)\,\sigma^2 \qquad (25)$$

A comparison between the values of Var ($R_n^*$) given by (25) and those computed from simulated samples is shown in Table I for the case $\sigma = 1.0$. It should be noted that for n approaching infinity, Var ($R_n^*$) given by (25) approaches $0.0741n.\sigma^2$, the asymptotic value given by Feller[3].

## Third and Fourth Moments of $M_n^*$ and $m_n^*$

The third and fourth moments of $M_2^*$ and $m_2^*$ can be shown to be given by

$$E(M_2^{*3}) = -E(m_2^{*3}) = (4\pi)^{-1/2}\,\sigma^3 \qquad (26)$$

and

$$E(M_2^{*4}) = E(m_2^{*4}) = 3\sigma^4\,/8. \qquad (27)$$

For n ≥ 3, the Monte Carlo method was used to compute the third and fourth moments of $M_n^*$. The computed values of these moments can be approximated by

$$\mu_3 = 0.03077\ (n-1)^{1.43572}.\sigma^3 \qquad (28)$$

and

$$\mu_4 = 0.05923\ (n-1)^{1.89966}.\sigma^4 \qquad (29)$$

where $\mu_3$ and $\mu_4$ denote the third and fourth central moments of $M_n^*$, respectively. The constants in (28) and (29) were obtained by means of the non-linear least squares method. It should be noted that (28) and (29) also apply to the third and fourth central moments of $-m_n^*$. A comparison between the simulated values and the values computed from the approximate formulas for these moments is made in Table II.

TABLE II: THIRD AND FOURTH CENTRAL MOMENTS OF $M_n^*$ FOR THE CASE OF INDEPENDENT NORMAL VARIABLES ($\sigma = 1.0$)

| n | $\mu_3$ Approximate formula | $\mu_3$ Simulated values | $\mu_4$ Approximate formula | $\mu_4$ Simulated values |
|---|---|---|---|---|
| 4 | 0.14899 | 0.14481 | 0.47743 | 0.47922 |
| 6 | 0.31023 | 0.32797 | 1.25994 | 1.23965 |
| 8 | 0.50290 | 0.54035 | 2.38750 | 2.39696 |
| 10 | 0.72142 | 0.73380 | 3.84841 | 3.83481 |
| 15 | 1.36044 | 1.38171 | 8.90832 | 8.89604 |
| 20 | 2.10908 | 2.09347 | 15.91265 | 15.86346 |
| 25 | 2.94956 | 2.94701 | 24.80150 | 24.83179 |
| 30 | 3.87039 | 3.86099 | 35.50378 | 35.52419 |
| 35 | 4.86336 | 4.87649 | 48.06576 | 48.07655 |
| 40 | 5.92222 | 5.90418 | 62.37755 | 62.36756 |
| 45 | 7.04204 | 7.03930 | 78.44182 | 78.46499 |
| 50 | 8.21882 | 8.20401 | 96.23752 | 96.19436 |

## Distribution of $M_n^*$

Having obtained the first four moments of $M_n^*$, the moments ratios $\beta_1$, $\beta_2$ and the Pearson criterion $\kappa$ can be computed fiom the equations

$$\beta_1 = \mu_3^2/\mu_2^3$$

$$\beta_2 = \mu_4/\mu_2^2$$

and

$$K = \frac{\beta_1 (\beta_2 + 3)^2}{4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1)}$$

where $\mu_r$, r = 2, 3, 4 are the rth central moments of the random variable.

The results are listed in Table III. For these values of $\beta_1$, $\beta_2$ and $\kappa$, the most appropriate frequency curve to be selected in approximating the distribution of $M_n^*$ is the Type I curve of the Pearson system[10,11]. The general equation of the Type I curve, also referred to as the beta distribution, is

TABLE III: NUMERICAL VALUES OF THE MOMENT RATIOS AND OF THE PEARSON CRITERION

| n | $\beta_1$ | $\beta_2$ | $\kappa$ |
|---|---|---|---|
| 5 | 0.34649 | 2.96955 | −0.25882 |
| 10 | 0.42126 | 3.34278 | −0.60521 |
| 15 | 0.44183 | 3.42851 | −0.78652 |
| 20 | 0.44789 | 3.44436 | −0.82205 |
| 25 | 0.44872 | 3.43680 | −0.79315 |
| 30 | 0.44735 | 3.42036 | −0.74523 |
| 35 | 0.44501 | 3.40050 | −0.69571 |
| 40 | 0.44225 | 3.37960 | −0.65032 |
| 45 | 0.43933 | 3.35879 | −0.61042 |
| 50 | 0.43638 | 3.33846 | −0.57559 |

$$f(x) = \frac{b^{1-p-q}}{B(p,q)} (a_1 + x)^{p-1} (a_2 - x)^{q-1}, \; a_1 \le x \le a_2 \qquad (30)$$

where $p, q$ = two parameters,
$a_1, a_2$ = two terminals,
$b = a_1 + a_2$,

and $B(p, q)$ is the beta function of p, q, defined by

$$B(p,q) = \int_0^1 t^{p-1} (1-t)^{q-1} \, dt$$

Among the three cases listed by Pearson and Johnson[12], only the first can apply. For this, the procedure of determining p, q, $a_1$ and $a_2$ is as follows. Letting

$$\tau = \frac{6(\beta_2 - \beta_1 + 1)}{3\beta_1 - 2\beta_2 + 6}$$

$$\varepsilon = \frac{\tau^2}{4 + \frac{1}{4}\beta_1 (\tau + 2)^2/(\tau + 1)}$$

it can be shown that $\tau = p+q$ and $\varepsilon = pq$, so that p and q are the roots of the quadratic

$$Z^2 - \tau Z + \varepsilon = 0 \qquad (31)$$

In this case $\mu_3 > 0$, p is then the smaller or the two roots. The value of b can be expressed[11] as

$$b = \tau[\mu_2 (\tau + 1)/\varepsilon]^{1/2} \qquad (32)$$

Locating the origin of the coordinates at the mode of the frequency curve allows $a_1$ and $a_2$ to be determined from

$$\frac{a_1}{p-1} = \frac{a_2}{q-1} = \frac{b}{p+q-2} = \frac{b}{\tau-2} \qquad (33)$$

The determination of p and q from (31) and $a_1$ and $a_2$ from (33) constitutes the process of fitting the distribution of $M_n^*$. The fitted curve can be located using the knowledge that the distance from the mode to the mean (expected value) is determined by

$$\delta = \frac{bp}{p+q} - a_1 \qquad (34)$$

For any value of n, the distance $\xi$ from the start of the curve to the mean, $bq/(p+q)$, is quarter than the expected value $\nu_1 = E(M_n^*)$ of $M_n^*$. This implies that the frequency curve starts at a point to the left of the initial or zero level (Fig. 2). The area under the curve between the start $-a_1$ and the initial level $X_o = \delta - \nu_1$ then represents the probability mass of the adjusted surplus at zero:

$$Pr = \text{Prob}\,(M_n^* = 0) = I_{x_o}(p,q) = \frac{1}{B(p,q)} \int_0^{x_o} t^{p-1} (1-t)^{q-1} \, dt \qquad (35)$$

where

$$x_o = \frac{a_1 + x_o}{b} \qquad (36)$$

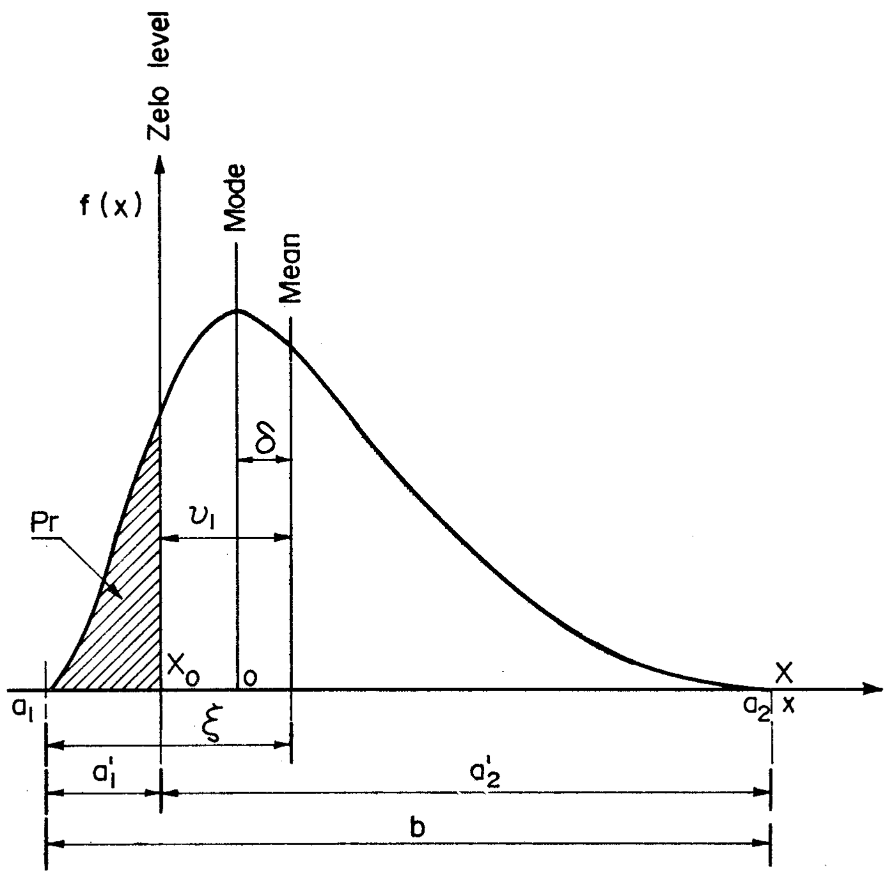


**Fig. 2.** Graphical representation of symbols used in describing the distribution of $M_n^*$,

The integral $I_{x_o}$ (p,q) is the incomplete beta function which was tabulated and first edited by Karl Pearson in 1954 (see Ref. 12). With the origin at the inital level, the probability density function of $M_n^*$ is deduced from (30) as

$$f(x) = C(a_1' + x)^{p-1} (a_2' - x)^{q-1}, \; 0 < x \ll a_2' \qquad (37)$$

in which

$$C = \frac{b^{1-p-q}}{B(p,q)}$$

$$a_1' = a_1 + x_o$$

$$a_2' = a_2 + x_o$$

At zero, $M_n^*$ has a probability mass given by (35). Table IV lists the values of the parameters and other related constants describing the distribution of $M_n^*$.

**TABLE IV:** CHARACTERISTICS OF THE DISTRIBUTION OF THE ADJUSTED SURPLUS FOR INDEPENDENT NORMAL VARIABLES ($\sigma = 1.0$)

| n | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|
| p | 3.84795 | 4.32210 | 4.09144 | 3.81221 | 3.57050 |
| q | 16.09100 | 22.00731 | 19.52161 | 16.66923 | 14.48116 |
| b | 12.01064 | 20.69107 | 23.53264 | 24.68230 | 25.39018 |
| $a_1$ | 1.90678 | 2.82529 | 3.36601 | 3.75575 | 4.06586 |
| $a_2$ | 10.10385 | 17.86577 | 20.16661 | 20.92653 | 21.32420 |
| $\delta$ | 0.41111 | 0.57124 | 0.71150 | 0.83836 | 0.95605 |
| C | $0.43544 \times 10^{-16}$ | $0.44539 \times 10^{-28}$ | $0.36867 \times 10^{-26}$ | $0.96788 \times 10^{-23}$ | $0.58324 \times 10^{-20}$ |
| $X_o$ | −0.98377 | −1.64657 | −2.13685 | −2.54132 | −2.89154 |
| $x_o$ | 0.07685 | 0.05697 | 0.05223 | 0.04920 | 0.04625 |
| Pr | 0.06430 | 0.03459 | 0.02441 | 0.01893 | 0.01535 |
| $a_1$ | 0.92301 | 1.17872 | 1.22916 | 1.21443 | 1.17431 |
| $a_2'$ | 11.08762 | 19.51233 | 22.30345 | 23.46783 | 24.21585 |

**TABLE V:** COMPUTED VALUES OF $\chi^2$

| n | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|
| $\chi^2$ | 23.682 | 16.722 | 16.117 | 25.189 | 31.872 |

Finally, a Chi-square test of goodness of fit was used. The computed values of the Chi-square statistic shown in Table V, indicate that fitting the distribution of $M_n^*$ by the Type I curve is acceptable.

## Distribution of $m_n^*$

For normal variables, the random variable $-m_n^*$ has the same distribution as

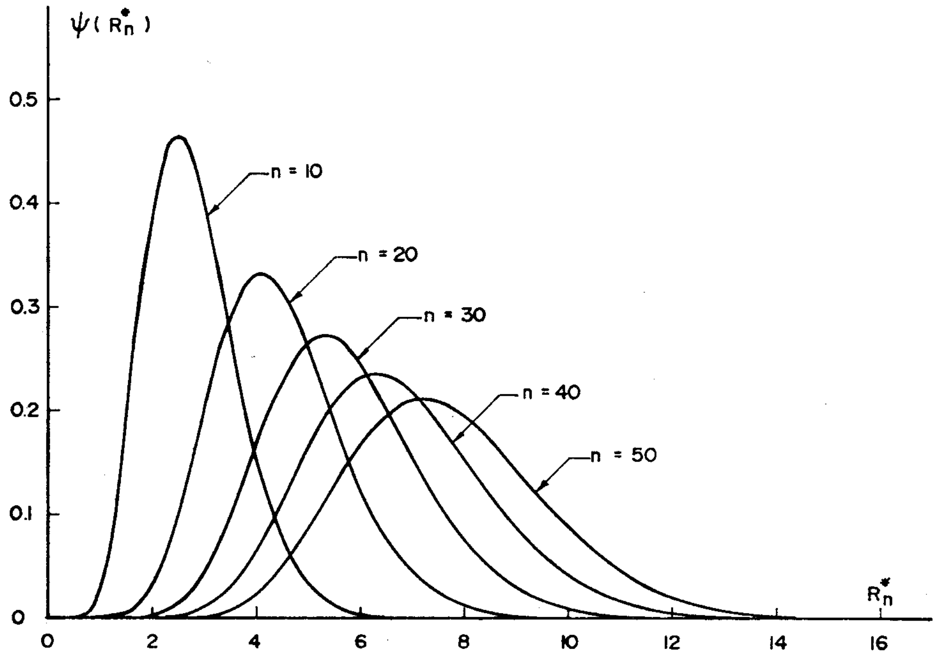


**Fig. 3.** Approximate probability density functions of $R_n^*$ for n = 10 (10) 50

$M_n^*$; consequently the distribution of $m_n$ can be readily deduced from that of $M_n^*$. Thus $m_n^*$ has a probability mass at zero and a probability density function g defined as follows:

$$\text{Prob } (m_n^*) = I_{x_0} (p,q)$$
$$g(m_n^*) = f(-m_n^*), \; -a_2' \leq m_n^* < 0$$

in which $I_{x_0}$ (p,q) and f are given by (35) and (37), respectively.

## Distribution of $R_n^*$

Following suggestion of Sutabutr[8], the Type III curve was used to fit the distribution of $R_n^*$. From the definition, $R_n^* \geq 0$, the two-parameter Pearson Type III curve, which is commonly known as the two-parameter gamma distribution, can be used. The probability density function of $R_n^*$ is then written as:

$$\psi(R_n^*) = \beta^{-\alpha}(R_n^*)^{\alpha-1}e^{-R_n^*/\beta}/\Gamma(\alpha), \; R_n^* \geq 0 \tag{38}$$

where $\Gamma(\alpha)$ denotes the gamma function of $\alpha$. The two parameters $\alpha$ and $\beta$ are determined from the expected value and variance of $R_n^*$ by

$$\alpha = [E(R_n^*)]^2 / \text{Var } (R_n^*)$$
$$\beta = \text{Var } (R_n^*) / E(R_n^*)$$

TABLE VI: RESULTS FROM FITTING THE DISTRIBUTION OF $R_n^*$

| n | $\alpha$ | $\beta$ | $\chi^2$ |
|---|---|---|---|
| 10 | 9.68608 | 0.28802 | 5.267 |
| 20 | 12.73857 | 0.34820 | 5.785 |
| 30 | 14.19927 | 0.40120 | 17.129 |
| 40 | 15.09630 | 0.44775 | 9.818 |
| 50 | 05.71841 | 0.48958 | 16.351 |

NOTES: Significance level = 0.05
Number of degrees of freedom = 29
Critical value = 42.557
Sample size used = 1,000

Numerical values of these parameters given in Table VI. Also shown in Table VI are the computed values of the Chi-squre statistic at various n. These values indicate the acceptability of the fitting. The approximate probability density functions of $R_n^*$ for n = 10(10)50 are shown in Fig. 3 for the case $\sigma = 1.0$.

## Summary and Conclusions

For a reservoir allowing neither spillage nor emptiness during its design life-time of n years where the outflows in different years are equal to the sample mean over n years of the annual inflows, the storage capacity was found to be optimally equal to the adjusted range. Under the above conditions, the adjusted partial sum, the adjusted surplus and the adjusted deficit represent the water content of the reservoir, the maximum amount of water in surplus, and the maximum amount of water in deficit, respectively. For the case where the inflows are distributed as independent normal variables, the water contents in different years were found to be distributed as normal variables; the maximum amount of water in surplus or in deficit was found to have a probability mass at zero and a probability density function elsewhere, which was fitted by the Type I curve; and the storage capacity was found to be fitted by the Type III curve.

## List of Symbols

The following symbols have been used in this paper:

$a_1, a_2$ = terminal of the beta distribution;
$a_1', a_2'$ = terminals of the distribution of the adjusted surplus;
b = distance between two terminals of the beta distribution;
B(p,q) = beta function of p and q;
C = constant in the distribution of the adjusted surplus;
Cov (. , .) = covariance;
E(.) = expected value;
f = probability density function of the adjusted surplus;
g = probability density function of the adjusted deficit;
$I_{x_o}$ (p,q) = incomplete beta function of p and q;
$m_n^*$ = adjusted deficit (maximum amount of water in deficit over n years);
$M_n^*$ = adjusted surplus (maximum amount of water in surplus over n years);
n = reservoir lifetime in years;
p,q = parameters of the beta distribution;
Pr = probability mass at zero;
$R_n^*$ = adjusted range (reservoir storage capacity);
$S_i$ = partial sum;
$S_i^*$ = adjusted partial sum (water content in the ith year);
$U_n^*$ = maximum of adjusted partial sums;
$V_i$ = positive part of $S_i^*$;
Var(.) = variance;
$W_i$ = negative part of $S_i^*$;
$X_k$ = annual inflow in year k;
$\overline{X}_n$ = sample mean;

| | | |
|---|---|---|
| $Y_k$ | = | net input in year k; |
| $\alpha$ | = | shape parameter of the gamma distribution; |
| $\beta$ | = | scale parameter of the gamma distribution; |
| $\beta_1$, $\beta_2$ | = | moment ratios; |
| $\Gamma(.)$ | = | gamma function; |
| $\delta$ | = | distance from the mode to the mean of the beta distribution; |
| $\varepsilon$ | = | pq; |
| $\kappa$ | = | Pearson criterion; |
| $\mu$ | = | mean of normal variables; |
| $\mu_r$ | = | central moment of order r; |
| $\psi$ | = | probability density function of $R_n^*$; |
| $\sigma$ | = | standard deviation of normal variables; |
| $\tau$ | = | p+q. |

**Acknowledgements**

This paper is based partly upon the Dissertation by H.N. Phien for the Degree of Doctor of Technical Science at the Asian Institute of Technology which was conducted under the supervision of the other two authors. The comments made by Professor Vujica Yevjevich of Colorado State University as external examiner are greatly appreciated.

**References**

1. Hurst, H.E. (1951) Longe-term Storage Capacity of Reservoirs. *Trans. Am. Soc. Civ. Eng.* **116,** 776–808.
2. Sen, Z. (1977) The Small Sample Expectation of Rescaled Population and Rescaled Adjusted Range. *Water Resour. Res.* **13,** 981–986.
3. Feller, W. (1951) The Asymptotic Distribution of the Range of Sums of Independent Variables. *Ann. Math. Stat.* **22,** 427–432.
4. Solari, M.E. and Anis, A.A. (1957) The Mean and Variance of the Maximum of the Adjusted Partial Sums of a Finite Number of Independent Normal Variates. *Ann. Math. Stat.* **28,** 706–716.
5. Salas-La Cruz, J.D. (1972) Range Analysis for Storage Problems of Periodic Stochastic Processes. *Hydrology Paper*, No. 57, Colorado State University, Fort Collins, Colorado.
6. Boes, D.C. and Salas-La Crus, J.D. (1973) The Expected Range and Expected Adjusted Range of Partial Sums of Exchangeable Random Variables. *J. Appl. Probab.* **10,** 671–677.
7. Salas-La Cruz, J.D. and Boes, D.C. (1974) Expected Range and Adjusted Range of Hydrologic Sequences. *Water Resour. Res.* **10,** 457–463.
8. Sutabutr, P. (1975) The Approximate Probability Density Functions of Range and Adjusted Range, Proceedings of the Second World Congress in Water Resources, New Delhi, Vol. 5, pp. 147–154.
9. Spitzer, F.A. (1956) A Combinatorial Lemma of Its Application to Probability Theory. *Trans. Am. Math. Soc.* **82,** 323–339.
10. Ord, J.K. (1972) *Families of Frequency Distributions*, Griffin, London.

11. Elderton, W.P. and Johnson, N.L. (1969) *Systems of Frequency Curves*, Cambridge University Press.
12. Pearson, E.S. and Johnson, N.L. (1968) *Tables of the Incomplete Beta Function*, Cambridge University press.

## บทคัดย่อ

ได้ศึกษาสภาพการเปลี่ยนแปลงปริมาณน้ำในอ่างเก็บน้ำในปีต่าง ๆ การเปลี่ยนแปลงของจำนวนสูงสุดของปริมาณน้ำที่เกินหรือต่ำกว่าระดับเริ่มต้น รวมทั้งการเปลี่ยนแปลงของปริมาณความสามารถในการเก็บกักภายในช่วงอายุการใช้งานของอ่างเก็บน้ำซึ่งถือว่าไม่มีการแห้งหรือล้นเลย การศึกษานี้มีข้อสมมุติฐานว่าปริมาณน้ำที่ไหลลงอ่างแต่ละปีมีการเปลี่ยนแปลงในแบบตัวแปรอิสระของการแจกแจงทางสถิติของ Gauss และปริมาณน้ำที่ไหลออกจากอ่างมีค่าเท่ากับค่าเฉลี่ยของตัวอย่างข้อมูลของปริมาณน้ำที่ไหลลงอ่าง

ได้พบว่าปริมาณน้ำที่เก็บไว้ในอ่างในปีต่าง ๆ เปลี่ยนแปลงตามการแจกแจงทางสถิติของ Gauss ส่วนปริมาณสูงสุดของน้ำที่ขาดและเกินนั้นมีค่า probability mass ที่ศูนย์ ส่วนที่อื่นมีค่าเป็น probability แบบต่อเนื่องซึ่งสามารถอธิบายโดยใช้เส้นโค้ง Pearson แบบที่ 1 สำหรับปริมาณความสามารถในการเก็บกักสามารถอธิบายโดยใช้เส้นโค้ง Pearson แบบที่ 3 การใช้เส้นโค้ง Pearson อธิบายนี้เราได้ประเมินความเหมาะสมโดยอาศัยการทดสอบทางสถิติซึ่งเรียกว่า Chi-square