# NOTES ON BIAS IN ESTIMATION

By M. H. QUENOUILLE

*Research Techniques Unit, London School of Economics and Political Science*

1. One of the commonest problems in statistics is, given a series of observations $x_1, x_2, \ldots, x_n$, to find a function of these, $t_n(x_1, x_2, \ldots, x_n)$, which should provide an estimate of an unknown parameter $\theta$.

The desirable properties of estimation procedures have been discussed fully elsewhere. They are:

(*a*) That the estimator should be efficient according to some definition of efficiency previously arranged. Most commonly, the reciprocal of the variance of the estimates is taken as a measure of its efficiency, as this is most useful where central limit theory may be relevant.

(*b*) That the estimator should utilize all the information contained in the observations, $x_1, x_2, \ldots, x_n$ concerning the parameter $\theta$. This is not always possible, but, if such an estimator exists, it is called sufficient.

(*c*) That the estimator should be consistent, i.e. $t_n$ converges in some probabilistic sense to $\theta$, usually $\lim_{n\to\infty} t_n \to \theta$.

(*d*) That the estimator should be unbiased, i.e. $E(t_n) = \theta$.

The method of maximum likelihood is popular in that it satisfies properties (*a*) to (*c*), whence, by evaluating $E(t_n)$, an unbiased statistic may be derived. That such evaluation is necessary is obvious when it is remembered that $\psi(t_n)$ is, by the same theory, the estimator of $\psi(\theta)$, and, since in general $E[\psi(t_n)] \neq \psi[E(t_n)]$, it will be the exception rather than the rule for a maximum-likelihood estimator to be unbiased.

Provided the exceptions may be simply evaluated no real difficulty arises. However, often the complexity of the evaluation presents a major drawback and some simple approach is then desirable.

2. If the observations are taken in random order, the estimator $t_n$ may often be written

$$t_n = t_n(k_1, k_2, \ldots, k_m),$$

where $k_1, k_2, \ldots, k_m$ are unbiased estimates of the cumulants $\kappa_1, \kappa_2, \ldots, \kappa_m$. Then, provided that

$$\left.\begin{array}{ll} (a) & m \text{ is independent of } n, \\ (b) & \text{the function } t_n \text{ is capable of Taylorian expansion}, \\ (c) & \text{all of the cumulants are finite}, \\ (d) & t_n \text{ is consistent, i.e. } \theta = \lim_{n\to\infty} t_n(k_1 \ldots k_m), \end{array}\right\} \qquad \text{(I)}$$

it follows that

$$t_n - \theta = \Sigma(k_i - \kappa_i)\left(\frac{\partial t_n}{\partial k_i}\right)_{k_i=\kappa_i} + \Sigma\Sigma(k_i - \kappa_i)(k_j - \kappa_j)\left(\frac{\partial^2 t_n}{\partial k_i \partial k_j}\right)_{k_i=\kappa_i} + \ldots.$$

Since the moments of the estimators, $k_i$, are power series in $1/n$, it follows that $E(t_n - \theta)$, i.e. the bias in $t_n$, is also expressible as a power series in $1/n$.

The conditions (I) are undoubtedly more stringent than they need be. For instance, the higher cumulants need not exist. Further, it appears likely that I (*b*) is a necessary condition if I (*d*) is to hold. However, the main point is that for a wide variety of statistics it is true that

$$E(t_n - \theta) = \frac{a_1}{n} + \frac{a_2}{n^2} + \frac{a_3}{n^3} + \ldots.$$

3. If this is so, and $t'_n = nt_n - (n-1)t_{n-1}$, then

$$E(t'_n) = \theta - \frac{a_2}{n^2} - \frac{a_2 + a_3}{n^3} - \ldots,$$

and hence $t'_n$ is biased to order $1/n^2$ only. Similarly, $t''_n = [n^2 t'_n - (n-1)^2 t'_{n-1}]/[n^2 - (n-1)^2]$ is biased to order $1/n^3$, and so on.

Alternatively, it is possible to use the statistics calculated from any subset of the observations to achieve corrections for bias. A further approach of particular interest occurs when $n = 2p$. Here, we may use $t'_{2p} = 2t_{2p} - t_p$ as being free from bias to order $1/n^2$.

Procedures such as these may supply approximate corrections for bias provided that efficiency of estimation is not lost in the process. To achieve this in general, it is necessary to use $\bar{t}_{n-1}$, the average of estimates from all possible sets of $n-1$ observations, instead of $t_{n-1}$, and similarly $\bar{t}_{n-2}$ instead of $t_{n-2}$, etc. With this provision, it appears likely that little, if any, loss of efficiency will result.

4. For instance, many of the statistics $t_n$ may be derived from an estimation procedure of form

$$\sum_{i=1}^{n} G(x_i, t_n) = 0.$$

The variance of $t_n$ to the first order may be estimated from this equation by a $\delta$ technique such as has been described by Weatherburn (1952, pp. 130 et seq.) and Kendall (1943, pp. 208 et seq.). In the simplest instance it is possible to represent the argument as follows. If $\mu = E(x_i)$, then

$$n\delta t_n = H(\theta) \sum_{i=1}^{n} \delta x_i,$$

where

$$H(\theta) = E\left[\frac{\partial}{\partial \mu} G(\mu, \theta)\right] \Big/ E\left[\frac{\partial}{\partial \theta} G(\mu, \theta)\right],$$

if both expectations exist. (If they do not exist, then generally the basic equation is changed to one of the form

$$n\delta t_n = \sum_{i=1}^{n} \delta H(x_i \theta),$$

to which a similar argument may be applied.)

Thus

$$\operatorname{var} t_n = \frac{[H(\theta)]^2}{n} \operatorname{var} x,$$

$$(n-1)\,\delta t_{n-1} = H(\theta) \sum_{i \neq j} \delta x_i,$$

$$(n-1)\,\delta \bar{t}_{n-1} = H(\theta) \frac{1}{n} \sum_{j=1}^{n} \sum_{i \neq j} \delta x_i$$

$$= H(\theta) \frac{n-1}{n} \sum_{i=1}^{n} \delta x_i.$$

Hence
$$\delta t_n' = n\delta t_n - (n-1)\,\delta \bar{t}_{n-1}$$
$$= \frac{H(\theta)}{n} \sum_{i=1}^{n} \delta x_i$$

and
$$\operatorname{var} t_n' = \frac{[H(\theta)]^2}{n} \operatorname{var} x = \operatorname{var} t_n$$

to order $1/n$.

This implies that this correction affects the standard error by a factor of $1/n$ at most, i.e.

$$\text{S.E. of } t_n' = (\text{S.E. of } t_n)\{1 + O(1/n)\}.$$

Since, in general, the standard error of $t_n$ will decrease with $n$ (usually proportional to $n^{-\frac{1}{2}}$), the correction will affect the dispersion of the distribution by $o(1/n)$ (usually, $O(n^{-\frac{3}{2}})$, i.e. by a small amount in comparison with the bias). The reduction in bias achieved by using $t_n$ is consequently not obtained at the expense of a comparable increase in the dispersion of the distribution of the estimator.

5. As a first illustration, suppose $\theta = \sigma^2$ for a normal distribution, and $t_n = \sum_{i=1}^{n} (x_i - \bar{x})^2/n$. Then

$$t_n' = nt_n - (n-1)\,\bar{t}_{n-1}$$
$$= \frac{n \sum\sum_{i<j} (x_i - x_j)^2}{n^2} - \frac{n-1}{n} \sum_{k=1}^{n} \frac{\overset{\text{Excluding } k}{\sum\sum_{i<j}} (x_i - x_j)^2}{(n-1)^2}$$
$$= \sum\sum_{i<j} (x_i - x_j)^2 \left[\frac{1}{n} - \frac{n-2}{n(n-1)}\right]$$
$$= \frac{1}{n(n-1)} \sum\sum_{i<j} (x_i - x_j)^2$$
$$= \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

and
$$\operatorname{var} t_n' = \left(\frac{n}{n-1}\right)^2 \operatorname{var} t_n.$$

Similarly, if $t_{2p}' = 2t_{2p} - \bar{t}_p$, then
$$t_{2p}' = \frac{1}{2p-1} \sum_{i=1}^{2p} (x_i - \bar{x})^2$$

and
$$\operatorname{var} t_{2p}' = \left(\frac{2p}{2p-1}\right)^2 \operatorname{var} t_{2p} = \frac{2\sigma^4}{2p-1}.$$

If, alternatively, $\bar{t}_p$ is calculated from only one pair of possible sets of $p$ observations, say $x_1$ to $x_p$ and $x_{p+1}$ to $x_{2p}$, then
$$t_{2p}' = \frac{1}{2p} \sum_{i=1}^{2p} x_i^2 - \frac{1}{p^2} \left(\sum_{i=1}^{p} x_i\right)\left(\sum_{i=p+1}^{2p} x_i\right)$$

and
$$\operatorname{var} t_{2p}' = \frac{p+1}{p^2} \sigma^4.$$

Thus averaging over only one pair of the possible sets results in a decrease in the efficiency of estimation of
$$\frac{(2p-1)(p+1)}{2p^2} - 1 = \frac{p-1}{2p^2}.$$

6. As a numerical illustration, suppose that it is desired to estimate $\theta = 1/\mu$ from a series of observations taken from a normal distribution. Then

$$t_n = n \Big/ \sum_{i=1}^{n} x_i$$

and

$$t'_n = \frac{n^2}{\sum_{i=1}^{n} x_i} - \frac{(n-1)^2}{n} \sum_{i=1}^{n} \frac{1}{\sum_{j \neq i} x_j},$$

$$\operatorname{var} t_n \sim \operatorname{var} t'_n \sim \frac{\sigma^2}{n\mu^4}.$$

The values in column 1 of Table 1 were random observations from a normal distribution with $\mu = 2$, $\sigma^2 = 1$ (using the first ten random numbers in Fisher & Yates's (1953) Statistical Tables).

Table 1

| $x_i$ | $18{\cdot}32 - x_i$ | $t_{n-1} = 9/(18{\cdot}32 - x_i)$ |
|---|---|---|
| 0·18 | 18·14 | 0·4961 |
| 4·00 | 14·32 | 0·6285 |
| 1·04 | 17·28 | 0·5208 |
| 0·85 | 17·47 | 0·5152 |
| 2·14 | 16·18 | 0·5562 |
| 1·01 | 17·31 | 0·5199 |
| 3·01 | 15·31 | 0·5879 |
| 2·33 | 15·99 | 0·5629 |
| 1·57 | 16·75 | 0·5373 |
| 2·19 | 16·13 | 0·5580 |
| 18·32 | $164{\cdot}88 = 9 \times 18{\cdot}32$ | 5·4828 |

Then

$$t_n = 1/1{\cdot}832 = 0{\cdot}54585,$$
$$t'_n = 5{\cdot}4585 - 9 \times 0{\cdot}54828 = 0{\cdot}5240.$$

Here, owing to the high value of $s^2$ ($= 1{\cdot}4$), this latter value has been corrected more than it would be using the exact formula

$$E(t_n) = \frac{n}{\sigma^2} \exp\left(-\frac{n\mu^2}{2\sigma^2}\right) \int_0^{\mu} \exp\left(\frac{n\mu^2}{2\sigma^2}\right) d\mu \quad \text{(see appendix)}$$

$$= \frac{1}{\mu} + \frac{\sigma^2}{n\mu^3} + \ldots \quad \text{for large } \mu.$$

This might be compared with

$$t_{n-1} = \frac{n-1}{T - x_i} = \frac{n}{T}\left[1 + \frac{n(x_i - \bar{x})}{(n-1)T} + \frac{n^2(x_i - \bar{x})^2}{(n-1)^2 T^2} + \ldots\right], \quad \text{where} \quad T = \sum_{i=1}^{n} x_i,$$

$$\bar{t}_{n-1} = \frac{n}{T} + \frac{n^2 s^2}{(n-1)T^3} + \ldots,$$

$$nt_n - (n-1)\bar{t}_{n-1} = \frac{n}{T} - \frac{n^2 s^2}{T^3} - \ldots = \frac{1}{\bar{x}} - \frac{s^2}{n\bar{x}^3} - \ldots.$$

In this instance it might be noted that the procedure will probably break down if $n\mu^2/\sigma^2$ is small. This should be apparent from the behaviour of the $t_{n-1}$, which will vary in sign.

7. Consider next an inverse sampling procedure. Suppose the proportion, $\pi$, of individuals with a given characteristic is to be estimated, and sampling continues until $r$ individuals with the characteristic are observed. Let $n$ be the total number of individuals.

Let $t_n = r/n$, then since the last individual is constrained to have the characteristic, there are only $n-1$ values of $t_{n-1}$ to be considered, and

$$\bar{t}_{n-1} = \frac{1}{n-1}\left[(r-1)\frac{r-1}{n-1} + (n-r)\frac{r}{n-1}\right] = \frac{nr-2r+1}{(n-1)^2}.$$

Thus

$$t'_n = n\frac{r}{n} - (n-1)\frac{nr-2r+1}{(n-1)^2} = \frac{r-1}{n-1},$$

which actually is strictly unbiased. Alternatively, if $1/\pi$ is to be estimated, then $t_n = n/r$, and

$$\bar{t}_{n-1} = \frac{1}{n-1}\left[(r-1)\frac{n-1}{r-1} + (n-r)\frac{n-1}{r}\right] = \frac{n}{r},$$

Thus $t'_n = t_n = n/r$, which again is strictly unbiased.

This indicates that the procedure may be useful in sequential estimation or inverse sampling.

8. It is possible to use simple extensions of this procedure to correct the bias in any combination of statistics, $f(t_n, u_m)$. The statistic

$$nmf(t_n, u_m) - (n-1)\,mf(t_{n-1}, u_m) - n(m-1)f(t_n, u_{m-1}) + (n-1)\,(m-1)\,f(t_{n-1}, u_{m-1})$$

is, for example, unbiased to order $1/n^2$ or $1/m^2$, whichever is the greater.

9. In general, where a series of concomitant observations, $y_1, y_2, \ldots, y_n$, are used in calculating $t_n$, both the bias and the efficiency will depend upon these observations, and therefore a simple correction of the above type will not be possible.

An important exception occurs when

$$t_n = \psi\left(\frac{\Sigma x_i \phi(y_i)}{\Sigma \phi^2(y_i)}\right). \tag{II}$$

In this instance, if

$$\xi = E[\Sigma x_i \phi(y_i)/\Sigma\phi^2(y_i)], \quad \sigma^2(t_n) = [\psi'(\xi)]^2 \operatorname{var} x / \sum_{i=1}^{n} \phi^2(y_i)$$

and

$$\frac{n-1}{\sigma^2(t_n)} = \Sigma \frac{1}{\sigma^2(t_{n-1})}.$$

Then, if $t'_n = nt_n - (n-1)\bar{t}_{n-1}$, where

$$\bar{t}_{n-1} = \left[\Sigma \frac{t_{n-1}}{\sigma^2(t_{n-1})} \Big/ \Sigma \frac{1}{\sigma^2(t_{n-1})}\right],$$

$t'_n$ is unbiased to order $1/n^2$, and $\operatorname{var} t_n = \operatorname{var} t'_n$ to order $1/n^2$. This formula may thus be put in the alternative forms

$$t'_n = nt_n - \Sigma \frac{\sigma^2(t_n)}{\sigma^2(t_{n-1})} t_{n-1}$$

$$= nt_n - \Sigma\left(1 - \frac{\phi^2(y_n)}{\Sigma\phi^2(y_i)}\right) t_{n-1}.$$

In a similar manner, if

$$t_n = \psi\left(\frac{\Sigma x_i\{\phi(y_i) - \overline{\phi(y_i)}\}}{\Sigma\{\phi(y_i) - \overline{\phi(y_i)}\}^2}\right),$$

then

$$\frac{n-2}{\sigma^2(t_n)} = \frac{n-1}{n}\Sigma\frac{1}{\sigma^2(t_{n-1})}$$

and

$$t'_n = (n-1)t_n - (n-2)\bar{t}_{n-1},$$

where

$$\bar{t}_{n-1} = \left[\Sigma\frac{t_{n-1}}{\sigma^2(t_{n-1})}\right] \Big/ \left[\Sigma\frac{1}{\sigma^2(t_{n-1})}\right]$$

is unbiased to order $1/n^2$.

It is also possible to split $2p$ observations into groups of $p$. Thus, if equation (II) holds,

$$\frac{1}{\sigma^2(t_{2p})} = \frac{1}{\sigma^2(t_{p,1})} + \frac{1}{\sigma^2(t_{p,2})},$$

and

$$t'_{2p} = 2t_{2p} - \bar{t}_p, \quad \text{where} \quad \bar{t}_p = \left(\Sigma\frac{t_p}{\sigma^2(t_p)}\right) \Big/ \left(\Sigma\frac{1}{\sigma^2(t_p)}\right),$$

is unbiased to order $1/n^2$ and has equal asymptotic efficiency to $t_{2p}$.

If, however, a correction is made for the mean, some loss in efficiency will arise from the difference between the values of $\overline{\phi(y_i)}$ for the two groups. If these are equal, the above approach may be used with no extra loss in efficiency, and, since there is no ordering involved in statistics of this type, approximate equality of these values may often be achieved by appropriate selection. For instance, if the $\phi(y_i)$ are ordered and the pairs of observations corresponding to alternate values are used in the two groups, then both $\Sigma\phi(y_i)$ and $\Sigma\phi^2(y_i)$ will be approximately equal for the two groups and $t'_{2p} = 2t_{2p} - \frac{1}{2}(t_{p1} + t_{p2})$ will frequently be a sufficiently accurate and unbiased statistic.

None of these results is, however, of much practical importance in that the bias may be calculated directly using $\psi''(\xi)\,\sigma^2(t_n)/2[\psi'(\xi)]^2$. Their interest lies in that they indicate that the same corrections applied to time-series statistics may be adequate for many purposes. This has already been suggested elsewhere (Quenouille, 1949).

10. Consider first the application of these corrections in the estimation of a serial covariance, say the $p$th. Here, if the mean is known, say $O$, an unbiased estimator exists:

$$t_n = (x_1x_{p+1} + x_2x_{p+2} + \ldots + x_{n-p}x_n)/(n-p),$$

with variance $\sigma^4/(n-p)$ in the case where no correlation exists in the series, and variance $\sigma^4A/(n-p)$ otherwise, where $A$ depends upon the correlations between the products in this expression.

There exist also $n$ estimators based upon $n-1$ observations. Denoting the estimator which omits the $i$th observation by $t_{n-1,i}$, this has $n-p-1$ terms if $i \leqslant p$ or $\geqslant n-p+1$, and $n-p-2$ terms otherwise. The variance of $t_{n-1,i}$ is correspondingly $\sigma^4/(n-p-1)$ or $\sigma^4/(n-p-2)$ when no correlation exists in the series, and $\sigma^4A_i/(n-p-1)$ or $\sigma^4A_i/(n-p-2)$ if correlation exists. Here, $A_i$ will differ for different $i$, but for large $n$ it will approximate to $A$ for all $i$.

If there is no serial correlation or if we ignore the differences in $A_i$, the analysis then gives

$$\begin{aligned} t_{n-1} &= \left(\Sigma\frac{t_{n-1}}{\sigma^2(t_{n-1,i})}\right) \Big/ \left(\Sigma\frac{1}{\sigma^2(t_{n-1,i})}\right) \\ &= (n-2)(n-p)t_n/[n(n-p-2)+2p] = t_n, \\ t'_n &= t_n. \end{aligned}$$

Thus the correction does not affect the estimate.

If the variation in $A_i$ is taken into account, a slightly different estimate is obtained. This is still unbiased, but is less efficient (though asymptotically of equal efficiency) probably as a consequence of the individual estimators not being fully efficient. For example, the products in $t_n$ are correlated with one another, and hence the end products contain information on $x_{n-p+1}x_{n+1}$, etc. The end-products should thus receive slightly greater weight for efficient estimation. Similarly, some of the products in $t_{n-1\,i}$ should receive greater weight. With these provisions, it appears likely that the above procedure would not lead to any loss in efficiency, i.e. the slight loss in efficiency (asymptotically zero) which occurs results from the use of inefficient estimates. This obviously is of no practical importance.

The effectiveness of corrections of this type in the general case is more difficult to prove. Their effectiveness might be demonstrated by considering the correction for the mean in the estimation of serial covariance.

If

$$t_n = \frac{1}{n-p}\left[\sum_{i=1}^{n-p} x_i x_{i+p} - \frac{\left(\sum_{i=1}^{n-p} x_i\right)\left(\sum_{i=1}^{n-p} x_{i+p}\right)}{n-p}\right],$$

then, when the $x_i$ are uncorrelated,

$$E(t_n) = -\frac{n-2p}{(n-p)^2}\sigma^2 = -\frac{\sigma^2}{n}\left[1+O\left(\frac{1}{n^2}\right)\right].$$

The expectations of the $t_{n-1}$ will vary. For instance, there will be a few terms of the type

$$E(t_{n-1,1}) = -\frac{E\left(\sum_{i=2}^{n-p} x_i\right)\left(\sum_{i=2}^{n-p} x_{i+p}\right)}{(n-p-1)^2} = -\frac{n-2p-1}{(n-p-1)^2}\sigma^2 = -\frac{\sigma^2}{n-1}\left[1+O\left(\frac{1}{n^2}\right)\right],$$

and

$$E(t_{n-1,2}) = -\frac{E\left(x_1+\sum_{i=3}^{n-p} x_i\right)\left(x_{p+1}+\sum_{i=3}^{n-p} x_{i+p}\right)}{(n-p-1)^2} = -\frac{\sigma^2}{n-1}\left[1+O\left(\frac{1}{n^2}\right)\right] \quad (p \neq 1),$$

but the majority ($n-4p$ out of $n$) will be of the form

$$E(t_{n-1,m}) = -\frac{E\left(\sum_{i=1}^{n-p} x_i - x_{m-p} - x_m\right)\left(\sum_{i=1}^{n-p} x_{i+p} - x_m - x_{m+p}\right)}{(n-p-2)^2}$$

$$= -\frac{n-2p-3}{(n-p-2)^2}\sigma^2 = -\frac{\sigma^2}{n-1}\left[1+O\left(\frac{1}{n^2}\right)\right].$$

Thus

$$E(\bar{t}_{n-1}) = -\frac{\sigma^2}{n-1}\left[1+O\left(\frac{1}{n^2}\right)\right] \quad \text{and} \quad E(t'_n) = O\left(\frac{1}{n^2}\right)$$

as required.

Similar results will hold for the serial correlation coefficients. It is, however, an open question as to whether the extra computation involved in calculating and using $\bar{t}_{n-1}$ is warranted compared with that involved in calculating and using $t_{(\frac{1}{2}n)}$. It appears likely that the use of the two half-series is sufficiently accurate for most practical purposes though this point requires further investigation.

## REFERENCES

FISHER, R. A. & YATES, F. (1953). *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh: Oliver and Boyd.
KENDALL, M. G. (1943). *The Advanced Theory of Statistics*, **1**. London: Griffin and Co.
QUENOUILLE, M. H. (1949). *J. R. Statist. Soc.* B, **11**, 68–84.
WEATHERBURN, C. E. (1952). *Mathematical Statistics.* Cambridge University Press.

## APPENDIX

### *Proof of a formula in* § 6

$$E(t_n) = \int_{-\infty}^{\infty} \frac{1}{x} \frac{\sqrt{n}}{\sigma\sqrt{(2\pi)}} \exp\left(-\frac{n(x-\mu)^2}{2\sigma^2}\right) dx$$

$$= \frac{\sqrt{n}}{\sigma} \int_{-\infty}^{\infty} \frac{1}{y\sqrt{(2\pi)}} \exp\{-\tfrac{1}{2}(y-a)^2\}\, dy,$$

where

$$a = \mu\sqrt{n}/\sigma,\ y = x\sqrt{n}/\sigma.$$

Let

$$I(a) = \int_{-\infty}^{\infty} \frac{1}{y\sqrt{(2\pi)}} \exp\{-\tfrac{1}{2}(y-a)^2\}\, dy,$$

then

$$\exp(\tfrac{1}{2}a^2)\, I(a) = \int_{-\infty}^{\infty} \frac{1}{y\sqrt{(2\pi)}} \exp(-\tfrac{1}{2}y^2 + ay)\, dy,$$

$$\frac{\partial}{\partial a}[\exp(\tfrac{1}{2}a^2)\, I(a)] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{(2\pi)}} \exp(-\tfrac{1}{2}y^2 + ay)\, dy$$

$$= \exp(\tfrac{1}{2}a^2).$$

Thus

$$\exp(\tfrac{1}{2}a^2)\, I(a) = \int_0^a \exp(\tfrac{1}{2}a^2)\, da,$$

the limits being determined by the fact that $I(a) = 0$ when $a = 0$.

Therefore

$$I(a) = \exp(-\tfrac{1}{2}a^2) \int_0^a \exp(\tfrac{1}{2}a^2)\, da$$

and

$$E(t_n) = \frac{n}{\sigma^2} \exp\left(-\frac{n\mu^2}{2\sigma^2}\right) \int_0^{\mu} \exp\left(\frac{n\mu^2}{2\sigma^2}\right) d\mu.$$