Estimating Mean Weight from Length Statistics

Biologists frequently need to estimate weights (y) of fish from their lengths (x), and numerous weight–length relationships have been developed empirically for different species, populations, life-history stages, seasons, and so on. Most or all take the form:

$$y = c \cdot x^n \quad \ldots\ldots(1)$$

where c is a constant and the exponent n commonly has values between 2.5 and 3.5, often being quite close to 3. From such relationships, weight can be estimated from length *for individual fish.*

Another common situation is where a quick estimate of the mean weight of the fish of a particular age-group of a species is desired — without going through the tedious computation of individual weights. For this purpose the weight corresponding to the observed mean length of the fish sampled has sometimes been used, but this estimate of course has a systematic bias that makes it too low (Ricker, 1958, p. 191). For any given mean length, this bias increases with the variability in length of the fish in the sample.

This note describes an unbiased and reasonably exact estimate of mean weight that can be obtained in any situation where the frequency distribution of lengths in a sample is approximately normal, where their mean and variance have been computed, and of course where a relationship of the type shown in equation 1 is available.

General theory. — Suppose x, the length of fish caught of a certain age-group in a species, is a normal random variable with mean μ and variance σ^2, i.e. x is $N(\mu, \sigma^2)$, so that its probability density is given by

$$\varphi(x) = (1/\sigma\sqrt{2\pi}) \exp\left[-(x-\mu)^2/2\sigma^2\right].$$

Then the expected value of a function of x, say $\psi(x)$, is given by

$$E[\psi(x)] = \int_{-\infty}^{\infty} \psi(x) \cdot \varphi(x)\, dx \quad \ldots\ldots(2).$$

Suppose, further, that the weight, y, of a fish is given by

$$y = c \cdot x^n$$
$$\text{i.e. } y = \psi(x) = c \cdot x^n$$

where c and n are *known.*

J. Fish. Res. Bd. Canada, 25(12): 2743–2747, 1968.
Printed in Canada.

The problem is to express the mean weight of a group of fish (usually all one age) in terms of the parameters (mean and variance) of the normal distribution of their lengths. Using equation 1 the mean weight can be expressed as

$$\mathrm{E}(y) = \int_{-\infty}^{\infty} c \cdot x^n \, \varphi(x) \, dx$$

$$= \int_{-\infty}^{\infty} (1/\sigma\sqrt{2\pi}) \cdot c \cdot x^n \exp\left[-(x-\mu)^2/2\sigma^2\right] dx \qquad \ldots.(3).$$

Case I. — For the special case of so-called isometric growth, $n = 3$, and we obtain from equation 3:

$$\mathrm{E}(y) = c/\sigma\sqrt{2\pi}\int_{-\infty}^{\infty} x^3 \exp\left[-(x-\mu)^2/2\sigma^2\right] dx$$

and by letting $t = (x-\mu)/\sigma$, we get

$$\mathrm{E}(y) = c/\sigma\sqrt{2\pi}\int_{-\infty}^{\infty} (\sigma t + \mu)^3 \exp(-t^2/2) \cdot \sigma \cdot dt$$

$$= c/\sqrt{2\pi}\int_{-\infty}^{\infty} (\sigma^3 t^3 + 3\sigma^2 t^2 \mu + 3\sigma t \mu^2 + \mu^3) \exp(-t^2/2)\, dt$$

$$= c(\sigma^3 \mathrm{I}_1 + 3\sigma^2\mu \mathrm{I}_2 + 3\sigma\mu^2 \mathrm{I}_3 + \mu^3 \mathrm{I}_4)$$

I_1 and I_3 are both zero, being respectively the first and third moments of a N(0, 1) variable, whereas I_2 is the second moment of a N(0, 1) variable and therefore equal to unity, and

$$\mathrm{I}_4 = 1/\sqrt{2\pi}\int_{-\infty}^{\infty} \exp(-t^2/2)\, dt$$

is simply the total cumulative distribution function of the N(0, 1) variable which, by definition, equals unity. We get therefore:

$$\mathrm{E}(y) = c \cdot (3\sigma^2\mu + \mu^3) = c \cdot \mathrm{E}(x^3).$$

The difference between $\mathrm{E}(y)$ and $c \cdot [\mathrm{E}(x)]^3$, in terms of the parameters of the distribution of x, is therefore

$$c \cdot (3\sigma^2\mu + \mu^3) - c \cdot \mu^3 = 3c \cdot \sigma^2\mu$$

or, in terms of the sample estimates

$$\mathrm{E}(y) - c \cdot [\mathrm{E}(x)]^3 = 3c \cdot s_x^2 \cdot \bar{x}.$$

The difference, as a fraction of the latter, i.e. of $c \cdot \bar{x}^3$, is $3 \cdot s_x^2/\bar{x}^2$.

In general, for integer values of n, it is easily shown that

$$\mathrm{E}(y) = c \cdot \left[\mu^n + \frac{n!}{(n-2)!2!}\mu^{n-2}\sigma^2 + \frac{3n!}{(n-4)!4!}\mu^{n-4}\sigma^4 + \right.$$

$$\left. + \frac{5n!}{(n-6)!6!}\mu^{n-6}\sigma^6 + ----\right] \qquad \ldots..(3)$$

which we abbreviate to

$$\mathrm{E}(y) = c.\ [\mu^n + a_1\mu^{n-2}\sigma^2 + a_2\mu^{n-4}\sigma^4 + a_3\mu^{n-6}\sigma^6 + ----] \qquad \ldots..(4)$$

where the number of terms is given by the integer part of $(n + 2)/2$. We get from equation 3 (omitting the multiplier c):

n	E(y)
1	μ
2	$\mu^2 + \mu^{\circ}\sigma^2$
3	$\mu^3 + 3\mu^1\sigma^2$
4	$\mu^4 + 6\mu^2\sigma^2 + 3\mu^{\circ}\sigma^4$
5	$\mu^5 + 10\mu^3\sigma^2 + 15\mu^1\sigma^4$
6	$\mu^6 + 15\mu^4\sigma^2 + 45\mu^2\sigma^4 + 5\mu^{\circ}9^6$

If the parameters of the distribution of x are now replaced by the sample statistics, it is possible to express $\mathrm{E}(y)$ in the appropriate number of terms of equation 4, and thus in terms of the distribution statistics of x.

As a simple illustration, assume x is N(100, 100) and $y = 0.5x^3$, with the sample values of x shown in Table I. We have:

Mean length	$= \bar{x} = 100$
Sample variance	$= s_x^2 = 133.33$
"True" mean weight (Table I)	$= \bar{y} = 518{,}252$
Estimated mean weight	$= c \cdot [\mathrm{E}(y)] = 0.5[(100)^3 + 3 \times 100 \times 133.33]$
	$= 520{,}000$
Weight of a fish of mean length	$= c[\mathrm{E}(x)]^3 = 0.5(100)^3$
	$= 500{,}000$

Thus the mean weight estimated from equation 4 is only 0.3% different from the mean weight of the sample, as computed for individual fish using equation 1. The weight of a fish of average length, however, is 3.8% less than the mean weight of the fish sampled.

TABLE I. Lengths of a sample of 16 fish from a normally distributed population, and the corresponding weights computed from $y = 0.5\,x^3$ and from $y = 0.5\,x^{3.3}$.

Length x	Weight $y = 0.5\,x^3$	Weight $y = 0.5\,x^{3.3}$
109	647,514	2,645,350
104	562,432	2,265,550
79	246,520	914,400
107	612,522	2,488,350
115	760,438	3,157,100
105	578,812	2,338,250
84	296,352	1,119,650
101	515,150	2,056,950
98	470,596	1,862,200
92	389,344	1,511,800
78	237,276	876,700
104	562,432	2,265,550
103	546,364	2,194,500
112	702,464	2,893,350
96	442,368	1,739,650
113	721,448	2,979,450
Sum 1600	8,292,032	33,308,800
Mean 100	518,252	2,081,800

Case II. — When n is not an integer, equation 4 may still be used in an appropriate form, the necessary a values being obtained by interpolation. Approximate values of a_1 and a_2 for noninteger $n \leq 5$ are given in Table II (obtained by nonlinear interpolation). In this case the number of terms to be used is the integral part of $(n + 3)/2$.

As an example, suppose that for the sample of 16 lengths in Table I, the weight–length relationship is:

$$y = 0.5x^{3.3}.$$

The corresponding computed weights are shown in column 3 of Table I.

From Table II, $a_1 = 3.77$ and $a_2 = 0.6$. Substituting these in equation 4 along with μ (taken as $\bar{x} = 100$) and σ^2 (taken as $s_x^2 = 133.33$), we obtain the following estimate of the mean weight of the fish in the sample:

$$\begin{aligned} c\cdot[\mathrm{E}(x^{3.3})] &= 0.5[(100)^{3.3} + 3.77 \times 100^{1.3} \times 133.33 + 0.6 \times 100^{-0.7} \times (133.33)^2] \\ &= 0.5[3{,}981{,}100 + 3.77 \times 398.11 \times 133.33 + 0.6 \times 0.39811 \times (133.33)^2] \\ &= 0.5[3{,}981{,}100 + 200{,}110 + 420] \\ &= 2{,}090{,}820. \end{aligned}$$

This is only 0.4% greater than the "true" mean weight of 2,081,800, as calculated for the individual fish using equation 1 (Table I).

TABLE II. Approximate values of the coefficients a_1 and a_2 of equation 4 in the text for values of n from 1.0 to 5.0.

n	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
						a_1					
1	0.00	0.08	0.15	0.23	0.32	0.42	0.52	0.63	0.74	0.87	1.00
2	1.00	1.14	1.31	1.48	1.66	1.86	2.07	2.29	2.52	2.75	3.00
3	3.00	3.25	3.51	3.77	4.04	4.32	4.62	4.95	5.29	5.64	6.00
4	6.00	6.36	6.73	7.11	7.50	7.90	8.31	8.72	9.14	9.57	10.00
						a_2					
3	0.00	0.2	0.4	0.6	0.9	1.2	1.5	1.8	2.2	2.6	3.0
4	3.00	3.5	4.1	4.9	5.9	7.1	8.4	9.7	11.3	13.0	15.0

The mean weight of a fish of average length in the sample is given by:

$$c[\mathrm{E}(x)]^{3.3} = 0.5(100)^{3.3} = 1{,}990{,}500.$$

This is 4.4% less than 2,081,800, so that again equation 4 gives a much better estimate of "true" mean weight.

Fisheries Research Board of Canada
Biological Station, Nanaimo, B.C.

L. V. PIENAAR
W. E. RICKER

Received for publication August 29, 1968.

REFERENCES

RICKER, W. E. 1958. Handbook of computations for biological statistics of fish populations. Bull. Fish. Res. Bd. Canada, No. 119.