ON MOMENTS OF ORDER STATISTICS FROM THE WEIBULL DISTRIBUTION

By Julius Lieblein

National Bureau of Standards

Summary. This note expresses the first two moments of the order statistics in samples from the Weibull distribution (sometimes referred to as the "third" asymptotic distribution of extreme values) in terms of known (incomplete B and Γ) functions. A similar procedure is applied to the "second" asymptotic distribution of extreme values.

1. Introduction. Explicit formulas in terms of a certain tabulated function were derived in an earlier paper [1] for the moments of order statistics for the "first" asymptotic distribution of largest values, with cdf

$$F(x) = \exp(-e^{-y}), \qquad y = (x - u)/\beta, \qquad -\infty < x < \infty. \tag{1.1}$$

This note extends the procedure to the other two asymptotic forms. However, the distributions for *smallest* rather than largest values are considered, with a view to possible application to breaking strength and fatigue problems. Without loss in generality for our purposes, the two other distributions may be taken (see [3]) as having cdf's

$$G(x) = \begin{cases} 1 - \exp[-(-x)^{-m}], & x \leqq 0, \\ 1, & x > 0; \end{cases} \tag{1.2}$$

$$H(x) = \begin{cases} 0, & x \leqq 0, \\ 1 - \exp(-x^{m}), & x > 0. \end{cases} \tag{1.3}$$

The parameter m is positive. These distributions are designated, respectively, the second and third asymptotic types (for *smallest* values). The distribution (1.3) has been applied extensively, mainly by Weibull [5], [6], [7], [8] and will be referred to by his name.

2. General approach. The main point of difficulty is in evaluating the double integrals that occur in the covariances, the means presenting little difficulty. It is interesting to proceed generally and try to discover other distributions for which the technique introduced in [1] will work.

If $P(x)$ is the cdf of the parent population, and $p(x)$ is the corresponding df, then the joint df of the ith and jth order statistics from $P(x)$ is [9]

$$p(x, y) = K[P(x)]^{i-1}[P(y) - P(x)]^{j-i-1}[1 - P(y)]^{n-j}p(x)p(y), \qquad -\infty < x \leqq y < \infty, \tag{2.1}$$

Received November 12, 1954.

where $K = K(n, i, j) = n!/[(i - 1)!\,(j - i - 1)!\,(n - j)!]$ and $x = x_i$ and $y = x_j$, with $i < j$.

If in (2.1) we write $P(y) - P(x) = [1 - P(x)] - [1 - P(y)]$, the covariances may be expressed in terms of integrals of the forms

$$\int_{-\infty}^{\infty}\int_{-\infty}^{y} xy\,[P(x)]^t\,[1 - P(y)]^u\,p(x)\,p(y)\,dx\,dy, \tag{2.2}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{y} xy\,[1 - P(x)]^t\,[P(y)]^u\,p(x)\,p(y)\,dx\,dy.$$

Leaving aside simple cases such as the rectangular distribution $P(x) = x$, parabolic distributions $P(x) = x^k$, and perhaps other simple forms for which the evaluation procedure in [1] is unnecessary, it will be seen on trial that success in using this procedure depends upon having $P(x)$ or $1 - P(x)$ of the form $e^{Q(x)}$, where Q is a reasonably simple function. Since $P(x)$ and $1 - P(x)$ are non-negative, this is always possible with real $Q(x)$.

The integral (2.2) is then a (finite) linear combination of integrals of the form

$$\psi(t, u) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} xye^{tQ(x)+uQ(y)}\,Q'(x)\,Q'(y)\,dx\,dy.$$

Here t and u are not necessarily the same as in (2.2).

To determine Q, we carry out the procedure as far as possible. The first step, integration by parts, gives

$$\psi(t, u) = \frac{1}{t}\int_{-\infty}^{\infty} ye^{(t+u)\,Q(y)}\,Q'(y)\,dy - \frac{1}{t}\int_{-\infty}^{\infty}\int_{-\infty}^{y} ye^{tQ(x)+uQ(y)}\,Q'(y)\,dx\,dy.$$

Calling the double integral $\psi_1(t, u)$, the second step, differentiating with respect to t, gives

$$\frac{\partial\psi_1(t, u)}{\partial t} = \int_{-\infty}^{\infty}\int_{-\infty}^{y} y\,Q(x)\,Q'(y)\,e^{tQ(x)+uQ(y)}\,dx\,dy.$$

This can be evaluated in two cases, (i) when the inner integral can be evaluated as it stands, and (ii) when the double integral is, apart from a constant factor, precisely of the form $\psi(t, u)$.

In case (i), we need to have $Q(x) = cQ'(x)$, whence $Q(x) = ae^{bx}$, which makes either $P(x)$ or $1 - P(x)$ doubly exponential. That is, we have the first asymptotic form for smallest values, $1 - \exp(-e^x)$. The simple transformation $x = -x'$ converts this into the distribution of largest values, $P(x) = \exp(-e^{-x})$, previously considered [1]. Thus this case has, essentially, already been treated.

Case (ii) requires that $xQ'(x) = cQ(x)$, so that $Q(x) = ax^b$ (which, of course, includes the exponential distribution). This gives, essentially, just the two forms (1.2) and (1.3).

It is rather interesting that the three asymptotic distributions of extreme values thus seemingly exhaust the possibilities for the method introduced in connection with one of them.

3. Results. The Weibull distribution (1.3) is of chief interest. The method in case (ii) applied to (2.1) and (2.2) for this distribution yields a simple first order differential equation in $\psi(t, u)$ whose solution is given by

$$t^{1+1/m}\psi(t, u) - u^{1+1/m}\psi(u, u) = \frac{1}{m^2}\Gamma(2 + 2/m)\int_u^t t^{1/m}(t + u)^{-2-2/m}\,dt.$$

The function $\psi(u, u)$ is readily evaluated and the integral may be expressed in terms of the B-function by the change of variable $t = uw\,/\,(1 - w)$. The final expression for $\psi(t, u)$ is

$$\psi(t, u) = m^{-2}(tu)^{-r}\,\Gamma(2r)\,B_p(r, r), \qquad r = 1 + 1/m; \quad p = t/(t + u); \quad m, t, u > 0. \tag{3.1}$$

For given n, the values of ψ are needed only for $t + u \leqq n$. To cut the calculations in half, we can use the relation

$$\begin{aligned}\psi(t, u) + \psi(u, t) &= m^{-2}(tu)^{-r}\,\Gamma(2r)\,[B_p(r, r) + B_{1-p}(r, r)]\\ &= m^{-2}(tu)^{-r}\Gamma^2(r).\end{aligned}$$

The above values enable us to write down the results for the Weibull distribution (1.3) as

$$\begin{aligned} E(x_i^k) &= \frac{n!}{(i-1)!\,(n-i)!}\Gamma\left(1 + \frac{k}{m}\right)\sum_{\mu=0}^{i-1}(-1)^{\mu}C_{\mu}^{i-1}\\ &\qquad\cdot(n + \mu - i + 1)^{-1-k/m},\\ &\qquad m > 0; \quad i = 1, 2, \cdots, n; \quad k = 0, 1, 2, \cdots;\\ E(x_i x_j) &= m^2K\sum_{\mu=0}^{i-1}\sum_{\nu=0}^{j-i-1}(-1)^{\mu+\nu}C_{\mu}^{i-1}C_{\nu}^{j-i-1}\psi(j - i + \mu - \nu,\\ &\qquad n - j + \nu + 1),\\ &\qquad m > 0; \quad i < j; \quad i, j = 1, 2, \cdots, n. \end{aligned} \tag{3.2}$$

Here K is as in (2.1), and the function ψ is given by (3.1).

For the second asymptotic distribution (1.2), the results are similar and may be obtained with little difficulty. The function corresponding to ψ is

$$\begin{aligned}\psi^*(t, u) &= m^{-2}(tu)^{-r'}[\Gamma^2(r') - \Gamma(2r')B_p(r', r')], \quad r' = 1 - 1/m,\\ &\qquad p = t/(t + u), \quad m > 2, \quad t,u > 0.\end{aligned} \tag{3.3}$$

The resulting moments are then

$$
(3.4)\quad
\begin{aligned}
E(x_i^k) &= \frac{n!}{(i-1)!\,(n-i)!}\,\Gamma\left(1-\frac{k}{m}\right)\sum_{\mu=0}^{i-1}(-1)^{\mu+k}C_\mu^{i-1} \\
&\qquad\cdot(n+\mu-i+1)^{-1+k/m} \\
&\qquad m>0;\ i=1,2,\cdots,n\quad k<m;\ k=0,1,2,\cdots; \\
E(x_i x_j) &= m^2K\sum_{\mu=0}^{i-1}\sum_{\nu=0}^{j-i-1}(-1)^{\mu-\nu}\,C_\mu^{i-1}C_\nu^{j-i-1} \\
&\qquad\cdot\psi^*(n-j+\nu+1,\ j-i+\mu-\nu), \\
&\qquad m>2;\ i<j;\ i,j,=1,2,\cdots,n.
\end{aligned}
$$

Here K is as in (2.1), and ψ^* is given by (3.3).

4. Remarks on application. The present results can be used in cases where m is known, perhaps from previous work. One important application of moments of order statistics is in finding minimum-variance unbiased estimators by means of linear functions of such statistics. An illustration of this for the case of the first asymptotic distribution will be found in [2], Appendix C. If m is not known, then there is available [4] a simple graphical procedure for obtaining an estimate which could then be used in the above formulas for first approximations.

REFERENCES

[1] J. Lieblein, "On the exact evaluation of the variances and covariances of order statistics in samples from the extreme-value distribution," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 282–287.

[2] J. Lieblein, *A New Method of Analyzing Extreme-value Data*, NACA Tech. Note, no. 3053, January 1954.

[3] *Probability Tables for the Analysis of Extreme-value Data*, National Bureau of Standards Applied Mathematics Series, no. 22 (1953), U. S. Government Printing Office, Washington, D. C., p. 2.

[4] H. C. S. Thom, "Frequency of maximum wind speeds," *Proc. Amer. Soc. Civil Engrs.*, Vol. 80 (1954), separate no. 539, 11 pp.

[5] W. Weibull, "A statistical theory of the strength of materials," *Ing. Vetenskaps Akad. Handl.*, Vol. 151 (Stockholm 1939), p. 15.

[6] W. Weibull, "The phenomenon of rupture in solids," *Ing. Vetenskaps Akad. Handl.*, Vol. 153 (Stockholm 1939), p. 17.

[7] W. Weibull, "A statistical distribution of wide applicability," *J. Appl. Mech.*, Vol. 18 (1951), pp. 293–297.

[8] W. Weibull, "Statistical design of fatigue experiments," *J. Appl. Mech.*, Vol. 19 (1952), pp. 109–113.

[9] S. S. Wilks, *Mathematical Statistics*, Princeton University Press, 1950, p. 90.