A MODIFIED EXTREME VALUE PROBLEM

By Benjamin Epstein[1]

Coal Research Laboratory, Carnegie Institute of Technology

1. Introduction and summary. Consider the following problem.

Particles are distributed over unit areas in such a way that the number of particles to be found in such areas is a random variable following the law of Poisson, with ν equal to the expected number of particles per unit area. Furthermore, the particles themselves are assumed to vary in magnitude according to a size distribution specified (independently of the particular unit area chosen) by a d.f. $F(x)$ defined over some interval $a \leq x \leq b$, with $F(a) = 0$ and $F(b) = 1$. The problem is to find the distribution of the smallest, largest, or more generally the nth smallest or nth largest particle in randomly chosen unit areas.

The problem as stated is not completely specified. To specify the distribution of smallest or largest particles in a unit area one must give a rule for dealing with those areas which contain no particles at all. More generally, in the case of the distribution of the nth smallest or nth largest particle, one must give a rule for dealing with those areas which contain $(n - 1)$ or fewer particles. There are at least two possible alternatives. One alternative is to omit none of the areas from consideration by setting up the following rule: if no particles are found in a given unit area then this area will be considered as one for which the smallest size particle is $x = b$ and for which the largest size particle is $x = a$. More generally, if $(n - 1)$ or fewer particles are found in a given unit area then this area will be considered as one for which the nth smallest size particle is $x = b$ and for which the nth largest size particle is $x = a$. A second alternative is to restrict attention to those areas which contain at least one particle (in the case of the distribution of smallest or largest values) or at least n particles (in the case of the distribution of the nth smallest or nth largest particle). In other words, this means finding the relevant conditional distribution.

From the point of view of the application of the theory of extreme values to fracture problems, there are some situations where the first model and other situations where the second model is the more appropriate in describing the phenomenon under investigation. In this paper section 2 will be devoted to a derivation of the distributions associated with the first alternative; in section 3 the conditional distributions will be described briefly.

2. The distributions under the first alternative. In this section we shall be concerned with the first alternative. To find the distribution of the nth smallest particle in unit areas, we first observe (the verification is left to the

---

[1] Present address, Department of Mathematics, Wayne University, Detroit, Michigan.

reader) that under the hypotheses of section 1, the number of particles having size $\leq x$ in a unit area is distributed according to the law of Poisson, with expected number equal to $\nu F(x)$. Next we note that the probability that the nth smallest particle in a unit area exceeds x in size is equal to the probability of finding exactly 0, or exactly 1, or exactly 2, $\cdots$, or exactly $(n - 1)$ particles of size $\leq x$ in that area. Therefore $G_n(x)$, the probability that the nth smallest size particle in a unit area is $\leq x$, is given by

$$G_n(x) = 1 - \sum_{j=0}^{n-1} e^{-\nu F(x)} \frac{(\nu F(x))^j}{j!}, \qquad x < b; \tag{1}$$

$$= 1, \qquad x \geq b,$$

where we have assigned to the size $x = b$ the probability $\Sigma_{j=0}^{n-1} e^{-\nu}(\nu^j/j)!$ which is just equal to the probability of finding fewer than n particles in a unit area.

If the d.f. $F(x)$ has a derivative $f(x)$ for all x lying in $a \leq x \leq b$, then $G_n(x)$ has a derivative for any value of $x \neq b$. Therefore the probability density for the nth smallest size particle is, for any $x \neq b$, given by the function $g_n(x)$ where

$$g_n(x) = e^{-\nu F(x)} \frac{(\nu F(x))^{n-1}}{(n-1)!} \nu f(x), \qquad a \leq x < b; \tag{2}$$

$$= 0, \qquad x < a, \qquad x > b.$$

A finite probability $\sum_{j=0}^{n-1} e^{-\nu} \frac{\nu^j}{j!}$ is assigned to $x = b$.

If one makes the transformation $y = \nu F(x)$ (for a similar transformation in extreme value theory see [1, page 371]), then (1), and (2) become

$$G_n^*(y) = 1 - \sum_{j=0}^{n-1} e^{-y} \frac{y^j}{j!}, \qquad y < \nu; \tag{1'}$$

$$= 1, \qquad y \geq \nu,$$

and

$$g_n^*(y) = \frac{e^{-y} y^{n-1}}{(n-1)!}, \qquad 0 \leq y < \nu; \tag{2'}$$

$$= 0, \qquad y < 0, \qquad y > \nu.$$

A finite probability $\sum_{j=0}^{n-1} e^{-\nu} \frac{\nu^j}{j!}$ is assigned to $y = \nu$.

The distribution of the smallest size particle in a randomly chosen area is found by letting $n = 1$ in equation 1.

In a similar way one can find the distribution of the nth largest particle in a randomly chosen unit area. $H_n(x)$, the probability that the nth largest size particle in a unit area is $\leq x$, is given by

$$
\begin{aligned}
H_n(x) &= 0, \qquad x < a; \\
(3) \qquad &= \sum_{j=0}^{n-1} e^{-\nu(1-F(x))} \frac{[\nu(1 - F(x))]^j}{j!}, \qquad x \geq a,
\end{aligned}
$$

where we have assigned to the size $x = a$ the probability $\sum_{j=0}^{n-1} e^{-\nu} \frac{\nu^j}{j!}$.

If, as before, $F(x)$ is assumed to have a derivative $f(x)$ for all x lying in $a \leq x \leq b$, then the probability density for the nth largest size particle is, for any $x \neq a$, given by the function $h_n(x)$ where

$$
\begin{aligned}
h_n(x) &= e^{-\nu(1-F(x))} \frac{[\nu(1 - F(x)]^{n-1}}{(n-1)!} \nu f(x), \qquad a < x \leq b; \\
(4) \qquad &= 0, \qquad x < a, \qquad x > b.
\end{aligned}
$$

A finite probability $\left|\sum_{j=0}^{n-1} e^{-\nu} \frac{\nu^j}{j!}\right.$ is assigned to $x = a$.

If one makes the transformation $z = \nu[1 - F(x)]$, then (3) and (4) become

$$
\begin{aligned}
H_n^*(z) &= 1 - \sum_{j=0}^{n-1} e^{-z} \frac{z^j}{j!}, \qquad z < \nu; \\
(3') \qquad &= 1, \qquad z \geq \nu,
\end{aligned}
$$

and

$$
\begin{aligned}
h_n^*(z) &= \frac{e^{-z} z^{n-1}}{(n-1)!}, \qquad 0 \leq z < \nu; \\
(4') \qquad &= 0, \qquad z < 0, \qquad z > \nu,
\end{aligned}
$$

with a finite probability $\sum_{j=0}^{n-1} e^{-\nu} \frac{\nu^j}{j!}$ assigned to $z = \nu$.

The distribution of the largest size particle in a randomly chosen unit area is found by letting $n = 1$ in equation 3.

3. Conditional distributions of the extreme values. The appropriate conditional distributions for the problem under consideration can be written down readily. The step function component which occurred in section 2 is no longer present since we restrict our attention only to those areas which contain at least n particles (in the general case of the distribution of nth smallest or nth largest size particles).

$G_n^c(x)$, the d.f. of the nth smallest particle in a unit area chosen at random from the class of areas containing at least n particles, is given by

$$
\begin{aligned}
G_n^c(x) &= 0, \qquad x < a; \\
(5) \qquad &= \frac{1 - \sum_{j=0}^{n-1} e^{-\nu F(x)} (\nu F(x))^j / j!}{1 - \sum_{j=0}^{n-1} e^{-\nu} \nu^j / j!}, \qquad a \leq x \leq b; \\
&= 1, \qquad x > b.
\end{aligned}
$$

Similarly $H_n^c(x)$, the d.f. of the nth largest particle in a unit area chosen at random from the class of areas containing at least n particles, is given by

$$
\begin{aligned}
H_n^c(x) &= 0, \qquad x < a; \\
(6) \qquad &= \frac{\sum_{j=0}^{n-1} e^{-\nu[1-F(x)]}[\nu(1 - F(x))]^j/j! - \sum_{j=0}^{n-1} e^{-\nu}\nu^j/j!}{1 - \sum_{j=0}^{n-1} e^{-\nu}\nu^j/j!}, \qquad a \leq x \leq b; \\
&= 1, \qquad x > b.
\end{aligned}
$$

4. General remarks and an application. It is interesting to note that the assumptions of section 1 lead to distribution functions in section 2 which are precisely the same as the asymptotic distributions of smallest, largest, or nth smallest, or nth largest values in samples of fixed size $N(N \to \infty)$ (see e.g. [1, p. 371]). In the problem treated in this paper, ν, the expected number of particles in a unit area, plays the role of N in the fixed sample size case, with the important difference that the distributions in the present paper are exact and not merely asymptotic.

The results of this paper have a direct bearing on certain aspects of fracture problems [2] and in particular on the dielectric breakdown of capacitors [3]. In the latter problem there appears to be ample justification for assuming that the breakdown voltage is influenced to a considerable degree by the presence of flaws known in the technical literature as conducting particles. These particles are spread individually and collectively at random throughout the area of the capacitor and, depending on their size, create a local weakening of the capacitor by reducing the nominal insulation thickness in the neighborhood of flaws. The voltage required to break down the capacitor is equal to that required to break it down at that spot where the greatest penetration has taken place.

In the dielectric problem the statistical distribution of largest values appropriate to the problem is given by (3) with $n = 1$, and the size distribution of conducting particles follows a law of the form $f(x) = \lambda e^{-\lambda x}$, $x > 0$. This is a situation where all the capacitors under test are part of the sample (since all must be tested to destruction) and those which happen to contain no defects (an event with probability $e^{-\nu}$) act as if the largest particle size is equal to $a = 0$. $e^{-\nu}$ simply represents the expected fraction of capacitors which have strength equal to the theoretical strength of the insulation.

The conditional distributions of section 3 would be more appropriate in the following sort of practical situation. Suppose that surface flaws spread at random on glass rods are known to reduce greatly the strength of the rods. Suppose that in a given sample of glass rods one takes out by some method of inspection those specimens which have no flaws. Then the strength distribution of the remaining specimens is a conditional distribution since each specimen must contain at least one flaw to be eligible as a member of the sample.

REFERENCES

[1] H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, 1946.
[2] B. EPSTEIN, "Statistical aspects of fracture problems," *J. Applied Phys.*, Vol. 19 (1948), pp. 140–147.
[3] B. EPSTEIN AND H. BROOKS, "The theory of extreme values and its implications in the study of the dielectric strength of paper capacitors," *J. Applied Physics*, Vol. 19 (1948), pp. 544–550.