Deriving the Gumbel Distribution

Michael P. Mersic

April 2, 2007

1 Historical Background

In 1928 the Extreme Value Distributions were first derived by Fisher and Tippett in *Limiting forms of the frequency distribution of the largest or smallest member of a sample.*, [1]. The result of Fisher and Tippett's paper is that these are the only Extreme Value Distributions:

$$\begin{aligned} G_\gamma(x) &= exp(-(1+\gamma x)^{\frac{-1}{\gamma}}) \qquad (1) \\ &\quad \text{where } 1+\gamma x > 0 \\ G_\gamma(x) &= exp(-e^{-x}) \qquad (2) \\ &\quad \text{where } \gamma = 0 \end{aligned}$$

Then in 1943 Gnedenko gave a more detailed proof of the above in *Sur la distribution limite du terme maximum d'une serie aleatoire* [2]. In 1958 E. J. Gumbel published his, now classic, text *Statistics of Extremes* [3]. His proof, however, was even less detailed then Fisher and Tippett's.

The case $\gamma = 0$ is often referred to as the Gumbel Distribution because of his extensive study of the function. I hope to provide a more detailed and basic derivation of the case $\gamma = 0$ then those derivations given in the above references. My presentation is based on the derivations given by Allen Gut in *Probability: A Graduate Course* [4] and by Samual Kotz in *Extreme Value Distributions: Theory and Applications* [5].

2 Deriving the Gumbel Distribution

The Gumbel Distribution is the extreme value distribution with $\gamma = 0$. In this paper I will derive the Gumbel Distribution. The derivation progresses in the following steps:

1. Define extreme value distribution.
2. Define max-stable distribution.
3. Prove extreme value = max-stable. See [4] for a proof.
4. Prove the extreme value distributions are of the form: $(G(x))^k = G(a_k x + b_k)$. This uses the fact that extreme value distributions are max-stable.
5. Prove when $a_k = 1$, $G(x) = exp(-e^{-x})$.

This presentation of the derivation of the Gumbel Distribution is based on the derivations given by Allan Gut in [4] and Samuel Kotz in [5].

2.1 Define extremal and max-stable

Definition 1. A non-degenerate distribution G is **extremal** if it can appear as a limit distribution of standardized partial maxima of iid random variables, that is, if there exist $\{c_n > 0, n \geq 1\}$, $\{d_n \in \Re, n \geq 1\}$ and a distribution function F such that,

$$(F(c_n x + d_n))^n \rightarrow G(x)$$

Definition 2. A non-degenerate distribution G is **max-stable** iff there exist $\{a_n > 0, n \geq 1\}$ and $\{b_n \in \Re, n \geq 1\}$, such that,

$$(G(a_n x + b_n))^n \rightarrow G(x)$$

2.2 Derive the Gumbel Distribution

Theorem 1. $G(x) = exp(-e^{-x})$ *is an extremal distribution.*

Proof. Suppose $G(x)$ is **extremal**. Then by definition of extremal distribution,

$$(F(c_n x + d_n))^n = G(x) \text{ as } n \rightarrow \infty \tag{3}$$

also,

$$(F(c_{nk}x + d_{nk}))^{nk} = G(x) \text{ for all } k \tag{4}$$

We also know that,

$$(F(c_n x + d_n))^{nk} = G(x)^k \text{ for all } k \tag{5}$$

and by the Convergence to Types theorem, there exist a_k and b_k such that,

$$(G(x))^k = G(a_k x + b_k) \tag{6}$$

Let $a_k = 1$, so $(G(x))^k = G(x+b_k)$. This is where the different extremal distributions are differentiated. With $a_k = 1$, we will derive $G(x) = exp(-e^{-x})$.

Now to solve for $b(k)$, recognize,

$$(G(x))^{kl} = G(x + b_k)^l = G(x + b_k + b_l) \tag{7}$$

and

$$(G(x))^{kl} = G(x + b_{kl}) \tag{8}$$

therefore

$$b(kl) = b(k) + b(l) \tag{9}$$

Since $b(k)$ is monotonically decreasing, the only solution to this Cauchy Functional Equation is (by Cauchy),

$$b(k) = \sigma log(k), \text{ where } \sigma \text{ is a constant.} \tag{10}$$

Let's go back to the beginning and plug $b(k)$ in,

$$(G(x))^k = G(x + \sigma log(k)) \tag{11}$$

take the log of both sides, and multiply by -1,

$$-log((G(x))^k) = -log(G(x + \sigma log(k))) \tag{12}$$

$$k(-logG((x))) = -log(G(x + \sigma log(k))) \tag{13}$$

take the log of both sides again,

$$log(k(-log(G((x)))) = log(-log(G(x + \sigma log(k)))) \tag{14}$$

$$log(k) + log(-log(G(x)) = log(-log(G(x + \sigma log(k)))) \tag{15}$$

Now, let $h(x) = log(-log(G(x)))$, so,

$$log(k) + h(x) = h(x - \sigma log(k)) \tag{16}$$

Let $k = e^{(\frac{x}{\sigma})}$ so that $x - \sigma log(k) = 0$, then

$$log(exp^{\frac{x}{\sigma}}) + h(x) = h(x - \sigma log(exp^{\frac{x}{\sigma}})) \tag{17}$$

$$h(x) = h(x - \sigma log(exp^{\frac{x}{\sigma}})) - log(exp^{\frac{x}{\sigma}}) \tag{18}$$

since $x - \sigma log(exp^{\frac{x}{\sigma}}) = 0$, we get

$$h(x) = h(0) - log(exp^{\frac{x}{\sigma}}) \tag{19}$$

$$h(x) = h(0) - \frac{x}{\sigma} \tag{20}$$

plug in the definition of h(x),

$$log(-log(G(x))) = h(0) - \frac{x}{\sigma} \tag{21}$$

exp both sides,

$$-log(G(x)) = exp[h(0) - \frac{x}{\sigma}] \tag{22}$$

that is,

$$-log(G(x)) = exp[-\frac{x - \sigma h(0)}{\sigma}] \tag{23}$$

Let $\mu = \sigma log(-log(G(0)))$,

$$G(x) = exp[-e^{\frac{-(x-\mu)}{\sigma}}] \tag{24}$$

This is $exp(-e^{-x})$ except for scaling. $\square$

References

[1] R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:190–190, 1928.

[2] B. Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Ann. Math.*, 44:423–453, 1943, Translated and reprinted in: Breakthroughs in Statistics, Vol. I, 1992, eds. S. Kotz and N.L. Johnson, Springer-Verlag, pp. 195-225.

[3] Emil Julius Gumbel. *Statistics of Extremes*. Columbia University Press, New York, 1958.

[4] Allan Gut. *Probability: A Graduate Course*. Springer, Springer Science+Business Media, inc., 233 Spring Street, New York, NY 10013, USA, 2005.

[5] Samuel Kotz and Saralees Nadarajah. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, 57 Shelton Street, Covent Garden, London WC2H9HE, 2000.