Notes on the Type I Extreme Value Distribution

John Kennan

If u is uniform on [0,1], then $y = -\log(u)$ has the unit exponential distribution, and $\zeta = -\log(y)$ has the type I extreme value distribution.

Suppose there are J alternatives, with payoffs $\tilde{v}_j = v_j + \zeta_j$, where $\{\zeta_j\}$ is a set of iid extreme value random variables. Define u_j by setting $\zeta_j = -\log(y_j)$ and $y_j = -\log(u_j)$. Then $\{u_j\}$ is a set of iid random variables that are uniformly distributed on [0,1].

Lemma 1

$$\Pr\left(\tilde{v}_1 = \max_j \tilde{v}_j\right) = \frac{e^{v_1}}{\sum_{j=1}^{J} e^{v_j}}$$

Proof:

Let d_j be an indicator for the event that $\tilde{v}_j$ is maximal. Then

$$\begin{aligned}\Pr\left(\zeta_j + v_j \le \zeta_1 + v_1 \mid u_1\right) &= \Pr\left(y_j \ge \Delta_j y_1\right) \\ &= \Pr\left(u_j \le (u_1)^{\Delta_j}\right) \\ &= (u_1)^{\Delta_j}\end{aligned}$$

where $\Delta_j = \exp(v_j - v_1)$. Thus

$$\begin{aligned}\Pr\left(d_1 = 1\right) &= \int \Pr\left(d_1 = 1 \mid u_1\right) du_1 \\ &\int \left(\prod_{j=2}^{J} \Pr\left(\zeta_j + v_j \le \zeta_1 + v_1 \mid u_1\right)\right) du_1 \\ &= \int_0^1 (u_1)^{A_1} du_1\end{aligned}$$

where $A_1 = \Sigma_{j>1} \Delta_j$. This yields

$$\Pr\left(d_1 = 1\right) = \frac{1}{1 + A_1} = \frac{e^{v_1}}{\sum_{j=1}^{J} e^{v_j}}$$

Lemma 2

$$E\left(\max_j \tilde{v}_j\right) = \gamma + \log\left(\sum_{j=1}^{J} e^{v_j}\right)$$

where γ is the Euler constant.

Proof:

$$\frac{\partial}{\partial v_k} E\left(\max_j \tilde{v}_j\right) = \Pr\left(d_k = 1\right)$$

$$= \frac{e^{v_k}}{\sum_{j=1}^{J} e^{v_j}}$$

Integrating this with respect to v_k gives

$$E\left(\max_j \tilde{v}_j\right) = c + \log\left(\sum_{j=1}^{J} e^{v_j}\right)$$

where c is constant with respect to v_k. In fact since the choice of k is arbitrary, c is in fact constant with respect to v_j for all j. Letting $v_j \to -\infty$ for $j > 1$ yields $c = E\zeta_1 = \gamma$.

Lemma 3

$$E\left(v_1 + \zeta_1 \mid d_1 = 1\right) = \gamma + \log\left(\sum_{j=1}^{J} e^{v_j}\right)$$

Proof:

Define the random variable X as

$$X = \max_{j>1}\left(u_j^{\frac{1}{\Delta_j}}\right)$$

so that $d_1 = 1$ iff $X \le u_1$ (almost surely). The distribution function of X is

$$F(x) = \prod_{j=2}^{J} \Pr\left(u_j \le x^{\Delta_j}\right) = x^{A_1}$$

So

$$\begin{aligned}\Pr\left(d_1 = 1\right) E\left(\zeta_1 \mid d_1 = 1\right) &= \int_0^1 \int_0^{u_1} -\log\left(-\log\left(u_1\right)\right) dF(x) du_1 \\ &= \int_0^1 -\log\left(-\log\left(u_1\right)\right)\left(u_1\right)^{A_1} du_1 \\ &= -\int_0^\infty \log\left(y_1\right)\exp\left(-A_1 y_1\right)\exp\left(-y_1\right) dy_1\end{aligned}$$

Define z = - log((1+A$_1$)y$_1$). Then

$$\begin{aligned}-\int_0^\infty \log\left(y_1\right)\exp\left(-A_1 y_1\right)\exp\left(-y_1\right) dy_1 &= \frac{1}{1+A_1}\int_{-\infty}^{\infty}\left(z + \log(1+A_1)\right) e^{-z} e^{-e^{-z}} dz \\ &= \frac{\gamma + \log(1+A_1)}{1+A_1}\end{aligned}$$

because exp(-z)exp(-exp(-z)) is the extreme value density function, and γ is the mean of the extreme value distribution. Thus

$$\frac{1}{1+A_1} E\left(\zeta_1 \mid d_1 = 1\right) = \frac{\gamma + \log(1+A_1)}{1+A_1}$$

and

$$\begin{aligned}E\left(\zeta_1 \mid d_1 = 1\right) &= \gamma + \log\left(1+A_1\right) \\ E\left(\zeta_1 + v_1 \mid d_1 = 1\right) &= \gamma + \log\left(e^{v_1} + e^{v_1} A_1\right) \\ &= \gamma + \log\left(\sum_{j=1}^{J} e^{v_j}\right)\end{aligned}$$