

12.25 The Largest of Many

WORKSHEET

The question

- (i) Let $\lambda > 0$. The independent random variables $X_1, X_2, \dots, X_n$ all have probability density function

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

and cumulative distribution function $F(x)$. The value of random variable Y is the largest of the values $X_1, X_2, \dots, X_n$. Show that the cumulative distribution function of Y is given, for $y \geq 0$, by

$$G(y) = (1 - e^{-\lambda y})^n.$$

- (ii) The values $L(\alpha)$ and $U(\alpha)$, where $0 < \alpha \leq \frac{1}{2}$, are such that

$$P(Y < L(\alpha)) = \alpha \quad \text{and} \quad P(Y > U(\alpha)) = \alpha.$$

Show that

$$L(\alpha) = -\frac{1}{\lambda} \ln(1 - \alpha^{\frac{1}{n}})$$

and write down a similar expression for $U(\alpha)$.

- (iii) Use the approximation $e^t \approx 1 + t$, for $|t|$ small, to show that, for sufficiently large n ,

$$\lambda L(\alpha) \approx \ln(n) - \ln\left(\ln \frac{1}{\alpha}\right).$$

- (iv) Hence show that the median of Y tends to infinity as n increases, but that the width of the interval $U(\alpha) - L(\alpha)$ tends to a value which is independent of n .
- (v) You are given that, for $|t|$ small, $\ln(1 + t) \approx t$ and that $e^3 \approx 20$. Show that, for sufficiently large n , there is an interval of width approximately $4\lambda^{-1}$ in which Y lies with probability 0.9.

Hints

Try the question on your own first. If you get stuck, read the hints *one at a time*, and go back to the question after each one. Hints marked *Pause and try* are the key steps.

? **Ask yourself** The largest of $X_1, \dots, X_n$ is at most y exactly when what happens?

STAGE 2 The quantiles

? **Ask yourself** Solve $G(L) = \alpha$ for L . Then solve $G(U) = 1 - \alpha$.

STAGE 3 Large n

? **Ask yourself** Write $\alpha^{1/n} = e^{(\ln \alpha)/n}$. For large n the exponent is small. What is $1 - \alpha^{1/n}$ approximately?

STAGE 5 A 90% interval

? **Ask yourself** Take $\alpha = 0.05$. What are $\ln \frac{1}{\alpha} = \ln 20$ and $\ln \frac{1}{0.95}$, approximately?

Try this next

- 1 The expected maximum.** Show that $E(Y) = \frac{1}{\lambda} \left(1 + \frac{1}{2} + \dots + \frac{1}{n}\right)$. (*Hint:* by memorylessness, after the first of the n lifetimes ends, the remaining $n - 1$ start afresh.)
- 2 The minimum.** Show that the smallest of $X_1, \dots, X_n$ is exponential with parameter $n\lambda$.
- 3 The median exactly.** Find the exact median of Y and compare it with the approximation $\frac{1}{\lambda}(\ln n - \ln \ln 2)$ for $n = 10$.
- 4 Uniform maxima.** If $X_1, \dots, X_n$ are uniform on $[0, 1]$, find the distribution of their maximum M , and show that $n(1 - M)$ is approximately exponential for large n . Which of the three extreme-value shapes is this?

Further reading

- **Stuart Coles, *An Introduction to Statistical Modeling of Extreme Values*** (Springer). The standard introduction, with real data on floods and sea levels.
- **Emil Julius Gumbel, *Statistics of Extremes*** (Dover). The classic.
- **David Spiegelhalter, *The Art of Uncertainty*** (Pelican), on rare events and how to think

about them.

The full guided solution, and what the question leads to at university, is in Section 12.25:
https://keithmcnulty.github.io/step_book_mcnulty/ch12/extremes.html