

STA 6351, Report.2.1

Carson Slater *Baylor University*

2.1

More Properties of Likelihood

A motivating example

Before we begin our more formal treatment of sufficiency, we consider a motivating example.

Example: Independent exponential random variables. Suppose X and Y are independent with common exponential density

$$f_{\theta}(x) = \theta e^{-\theta x} I_{[0, \infty)}(x)$$

where $\theta > 0$; that is, $\Theta = \mathbb{R}^+$. Let $U \sim U(0, 1)$ be independent of X and Y . Define $T = X + Y$. Finally, define the complementary fractions of that sum as

$$\tilde{X} = UT \quad \text{and} \quad \tilde{Y} = (1 - U)T.$$

In a reliability context, the randomization can be viewed as a randomized reassembly of component lifetimes. Suppose X and Y represent the independent lifetimes (in hours, say) of two identical components, each having an exponential lifetime distribution with rate parameter θ . If the components are used sequentially, the total operating time until both fail is

$$T = X + Y,$$

so that $T \sim \text{Gamma}(2, \theta)$. If, in an experiment or field setting, only the total service time T is observed, but not the individual component lifetimes, then the bivariate random variable

$$(\tilde{X}, \tilde{Y}|T) = (UT, (1 - U)T|T),$$

provides a stochastic model for the unobserved component lifetimes given the total. Here U serves as an allocation fraction, independent of T , expressing complete ignorance about how the total time was shared between the two components. The pair (\tilde{X}, \tilde{Y}) can thus be interpreted as a randomized reassembly of plausible component lifetimes consistent with the observed total T .

It will prove interesting to compare the joint distribution of X and Y with that of \tilde{X} and \tilde{Y} . To this end, we shall need the density of T . Since X and Y are independent,

$$\begin{aligned} Pr(T \leq t|Y = y) &= Pr(X + Y \leq t|Y = y) \\ &= \mathbb{E}[I\{X + Y \leq t\}|Y = y] \\ &= \int I\{x + y \leq t\} f_{\theta}(x) dx \end{aligned}$$

$$\begin{aligned}
&= \int I\{x \leq t - y\} f_\theta(x) dx \\
&= \int_{-\infty}^{t-y} f_\theta(x) dx \\
&= F_X(t - y) \quad (\text{CDF of } X \text{ evaluated at } t - y)
\end{aligned}$$

So, conditioning on the random variable Y instead of its value, we have

$$Pr(T \leq t|Y) = F_X(t - Y), \quad (1)$$

and therefore

$$\begin{aligned}
F_T(t) &= Pr(T \leq t) \\
&= \mathbb{E}Pr(T \leq t|Y) \\
&= \mathbb{E}[F_X(t - Y)]
\end{aligned} \quad (2)$$

Here, the first line is by definition, the second line makes use of the double expectation formula, and (2.1.2) is from (2.1.1). Applied to the present problem, we have (the random variable)

$$F_X(t - Y) = 1 - e^{-\theta(t-Y)} I_{\{Y \leq t\}}(Y, t). \quad (3)$$

With probability 1, $T \geq 0$. So, starting with (2.1.2) we substitute (2.1.3) and average out Y :

We can obtain the density for T by taking derivatives:

Now, since T and U are independent they have joint density

$$p_\theta(t, u) = t\theta^2 e^{-\theta t} I_{[0, \infty)}(t) I_{(0, 1)}(u).$$

Hence, for any set $B \in \mathbb{R}^+ \times \mathbb{R}^+$

$$Pr((\tilde{X}, \tilde{Y}) \in \tilde{B}) = \iint I_B(ut, (1-u)t) p_\theta(t, u) du dt. \quad (4)$$

Changing variables to $x = ut$, $du = dx/t$ in the inner integral, and reversing the order of integration using Fubini's theorem, we have

$$Pr((\tilde{X}, \tilde{Y}) \in B) = \iint I_B(x, t-x) t^{-1} p_\theta(t, x/t) dt dx.$$

Another change of variables in the inner integral, taking $y = t - x$, yields $\mathbf{t} = \mathbf{x} + \mathbf{y}$ and $d\mathbf{t} = d\mathbf{y}$. The region of integration $\mathbf{t} \in [x, \infty)$ maps to $\mathbf{y} \in [0, \infty)$. The Jacobian for this transformation is $|\frac{\partial \mathbf{t}}{\partial \mathbf{y}}| = 1$.

$$Pr((\tilde{X}, \tilde{Y}) \in B) = \iint I_B(x, y) (x+y)^{-1} p_\theta\left(x+y, \frac{x}{x+y}\right) dy dx.$$

It follows that \tilde{X} and \tilde{Y} have joint density

$$f_{\tilde{X}, \tilde{Y}}(x, y) = \frac{p_\theta\left(x+y, \frac{x}{x+y}\right)}{x+y}$$

Derivation of (2.1.5): Substituting $t = x + y$ and $u = x/(x + y)$ into $p_\theta(t, u) = t\theta^2 e^{-\theta t} I_{[0, \infty)}(t) I_{(0, 1)}(u)$:

$$p_\theta \left(x + y, \frac{x}{x + y} \right) = (x + y)\theta^2 e^{-\theta(x+y)} I_{[0, \infty)}(x + y) I_{(0, 1)} \left(\frac{x}{x + y} \right)$$

Therefore,

$$f_{\tilde{X}, \tilde{Y}}(x, y) = \frac{1}{x + y} \left[(x + y)\theta^2 e^{-\theta(x+y)} I_{[0, \infty)}(x + y) I_{(0, 1)} \left(\frac{x}{x + y} \right) \right] I_{[0, \infty)}(x) I_{[0, \infty)}(y)$$

For $x > 0$ and $y > 0$, we have $x + y > 0$ and $0 < x/(x + y) < 1$, simplifying the indicator functions:

$$f_{\tilde{X}, \tilde{Y}}(x, y) = \theta^2 e^{-\theta(x+y)} I_{[0, \infty)}(x) I_{[0, \infty)}(y). \quad (5)$$

Derivation of (2.1.6): The joint density of X and Y , being independent and $X, Y \sim f_\theta$, is $f_{X, Y}(x, y) = f_\theta(x)f_\theta(y)$.

$$\begin{aligned} f_\theta(x)f_\theta(y) &= \left(\theta e^{-\theta x} I_{[0, \infty)}(x) \right) \left(\theta e^{-\theta y} I_{[0, \infty)}(y) \right) \\ f_\theta(x)f_\theta(y) &= \theta^2 e^{-\theta(x+y)} I_{[0, \infty)}(x) I_{[0, \infty)}(y). \end{aligned} \quad (6)$$

So, the joint distribution of \tilde{X} and \tilde{Y} is the same as that for X and Y .

- That means that estimating θ with observed values of \tilde{X} and \tilde{Y} would be equivalent to using observed values of X and Y .
- However, note that I can compute values of \tilde{X} and \tilde{Y} by generating a value of T and a value of U and the latter does not depend on θ !
- Hence, $T = X + Y$ by itself provides as much information about θ as the pair (X, Y) .
- The sum T is called **sufficient** for the family of densities $\{f_\theta, \theta \in \Theta\}$.

This construction also illustrates an important theoretical point. Recall that a statistic $A = A(\mathbf{X})$ is said to be **ancillary** if its distribution does not depend on the underlying parameter; such statistics capture variation unrelated to θ . In the original variables, $T = X + Y$ and $V = X/(X + Y)$ form a **sufficient ancillary pair**. Because $T \sim \text{Gamma}(2, \theta)$ depends on the rate parameter θ while $V \sim \text{Uniform}(0, 1)$ does not, the factorization theorem implies that all information about θ is contained in T , and none in V .

The randomization step

$$(\tilde{X}, \tilde{Y}) = (UT, (1 - U)T), \quad \text{with } U \perp T,$$

preserves this structure: T carries the inferential content about θ , whereas the independent, uniformly distributed allocation fraction U represents purely ancillary variation. In this sense, the pair (\tilde{X}, \tilde{Y}) provides an explicit example of how observed data may be decomposed into a sufficient statistic (T) and an ancillary component (U), illustrating the independence principle that underlies the exponential-family framework.