# Bayesian statistics

## Lecture 2

**Dr. Rianne de Heide, February 15, 2024**

# Bruno de Finetti - 1974 - Theory of Probability

“My thesis, paradoxically, and a little provocatively, but nonetheless genuinely, is simply this:

PROBABILITY DOES NOT EXIST.

The abandonment of superstitious beliefs about the existence of Phlogiston, the Cosmic Ether, Absolute Space and Time,…, or Fairies and Witches, was an essential step along the road to scientific thinking. Probability, too, if regarded as something endowed with some kind of objective existence, is no less a misleading misconception, an illusory attempt to exteriorize or materialize our true probabilistic beliefs.”

# Harold Jeffreys 1961 - Theory of Probability p.viii

“Adherents of frequency definitions of probability have naturally objected to the whole system. But they carefully avoid mentioning my criticisms of frequency definitions, which any competent mathematician can see to be unanswerable. In this way they contrive to present me as an intruder into a field where everything was already satisfactory. I speak from experience in saying that students have no difficulty in following my system if they have not already spent several years in trying to convince themselves that they understand frequency theories.”

# Harold Jeffreys 1961 - Theory of Probability, p 38

“This distinction shows that theoretically a probability should always be worked out completely. We have again an illustration from pure mathematics. What is the 10,000th figure in the expansion of e? Nobody knows; but that does not say that the probability that it is a 5 is 0.1. By following the rules of pure mathematics we could determine it definitely, and the statement is either entailed by the rules or contradicted; in probability language, on the data of pure mathematics it is either a certainty or an impossibility.”

# Exercise 1

The number of offspring $X$ in a certain population has probability function $p(x \mid \alpha, \beta) = \alpha$ if $x = 0$ and $p(x \mid \alpha, \beta) = (1 - \alpha)\beta(1 - \beta)^{x-1}$ if $x = 1,2,\ldots$

where $\alpha$ and $\beta$ are unknown parameers lying in the unit interval.

Obtain the likelihood function when $r$ zeros and $n - r$ non-zero values $x_1, x_2, \ldots, x_{n-r}$ are obtained from $n$ independent observations on $X$.

Suppose $\alpha$ and $\beta$ have independent prior beta densities with parameters a an b and p en q respectively. Show that $\alpha$ and $\beta$ have independent posterior beta distributions and identify the posterior parameters.

# Exercise 2

Five machines are put on test for at most one hundred hours. Three of them fail at 65, 89 and 95 hours. The remaining machines are still working at one hundred hours. Assuming the lifetime of the machines has an exponential distribution with mean $\theta^{-1}$ derive the likelihood function. [Hint: What is the probability that a machine is still working at 100 hours?]

If the prior distribution for $\theta$ is a Gamma distribution $p(\theta) = GA(2,180)$, derive the posterior distribution of $\theta$.

Find the posterior mean and variance of $\theta$.

# Exercise 3

The inverted Pareto distribution has pdf $p(x\,|\,\theta, \delta) = \theta\delta^{\theta}x^{\theta-1}$ for $0 < x < \delta^{-1}, \theta > 0\ \delta > 0$ and zero otherwise. We will denote this by IP($\theta, \delta$).

Suppose $\delta = 2$. The prior for $\delta$ is taken as the improper density $p(\delta) \propto \delta^{-1}$. A random sample $x^n$ is taken from the IP($2$,$\delta$) distribution. Show that the density of the posterior distribution is

$$p(\delta\,|\,x^n) = \frac{2n\delta^{2n-1}}{t^{2n}},\ 0 < \delta < t, \text{ where } t = \min x_i^{-1}$$

# Exercise 4

Consider again the inverted Pareto distribution with pdf $p(x \mid \theta, \delta) = \theta\delta^{\theta}x^{\theta-1}$ for $0 < x < \delta^{-1}$, $\theta > 0$ $\delta > 0$ and zero otherwise. We will denote this by IP($\theta, \delta$).

Show that $p(\delta) \propto \delta^{-1}$ is Jeffreys' prior.

# Exercise 5

Find Jeffreys prior when the likelihood is Bin($n, \theta$), i.e.

$$f(x \mid n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

where n is a constant.

Make a drawing of Jeffreys prior for this model.