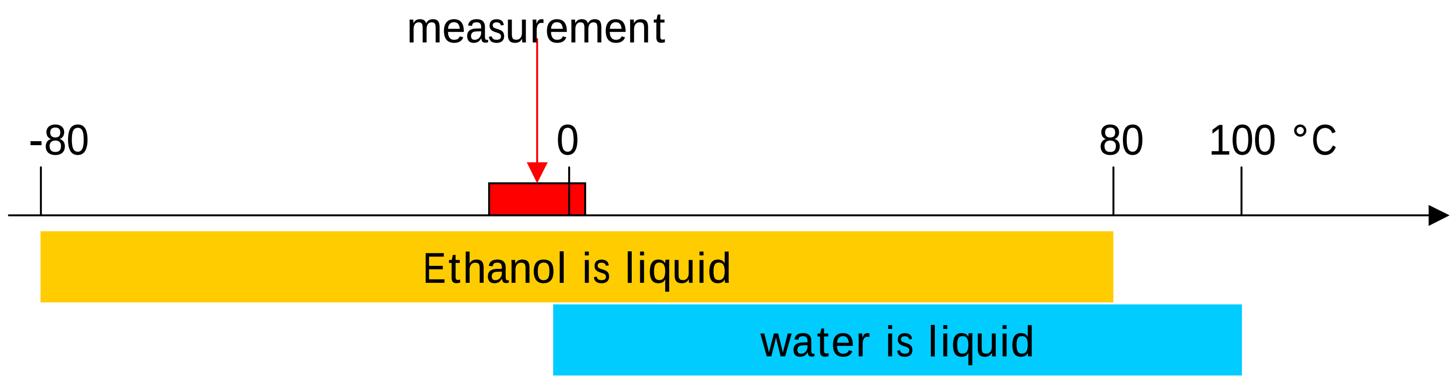
Bayesian hypothesis testing

- Example: A spacecraft is sent to a moon of Saturn and, using a penetrating probe, detects a liquid sea deep under the surface at 1 atmosphere pressure and a temperature of -3°C. However, the thermometer has a fault, so that the temperature reading may differ from the true temperature by as much as ±5°C, with a uniform probability within this range.

  Determine the temperature of the liquid, assuming it is water (liquid within $0<T<100$°C) and then assuming it is ethanol (liquid within $-80<T<80$°C). What are the odds of it being ethanol?

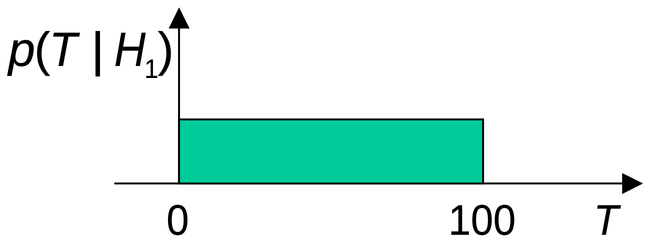


[based loosely on a problem by John Skilling]

Bayesian hypothesis testing

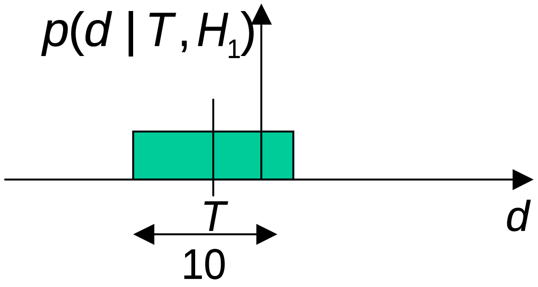
- Call the water hypothesis H_1 and the ethanol hypothesis H_2.
  For H_1:
  The prior on the temperature is

$$p(T \mid H_1) = \begin{cases} 0.01 & \text{for } 0 < T < 100 \\ 0 & \text{otherwise} \end{cases}$$

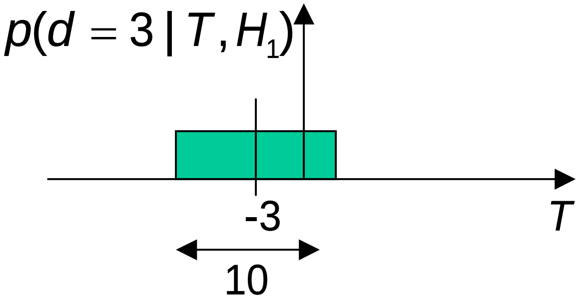


the likelihood of the temperature is the probability of the data *d*, given the temperature:

$$p(d \mid T, H_1) = \begin{cases} 0.1 & \text{for } |d - T| < 5 \\ 0 & \text{otherwise} \end{cases}$$



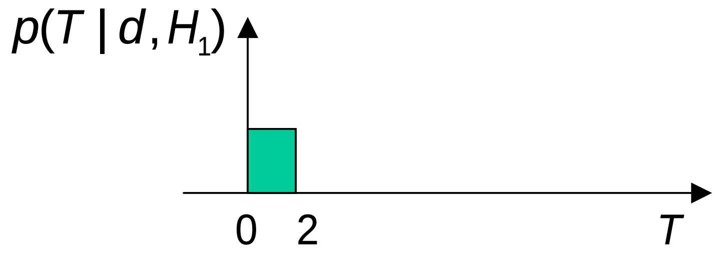
thought of as a function of T, for d=-3,



Bayesian hypothesis testing

The posterior for T is the normalised product of the prior and the likelihood, giving

$$p(T \mid d, H_1) = \begin{cases} 0.5 & \text{for } 0 < T < 2 \\ 0 & \text{otherwise} \end{cases}$$



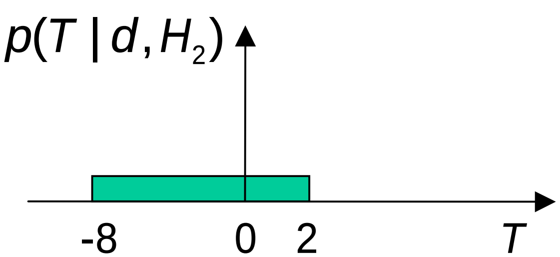
H_1 only allows the temperature to be between 0 and 2°C.

The *evidence* for water (as we defined it) is

$$p(d \mid H_1) = \int p(d \mid T, H_1) p(T \mid H_1) \mathrm{d}T = 0.002$$

- For H_2:
  By the same arguments

$$p(T \mid d, H_2) = \begin{cases} 0.1 & \text{for } -8 < T < 2 \\ 0 & \text{otherwise} \end{cases}$$



and the evidence for ethanol is

$$p(d \mid H_2) = \int p(d \mid T, H_2) p(T \mid H_2) \mathrm{d}T = 0.00625$$

Bayesian hypothesis testing

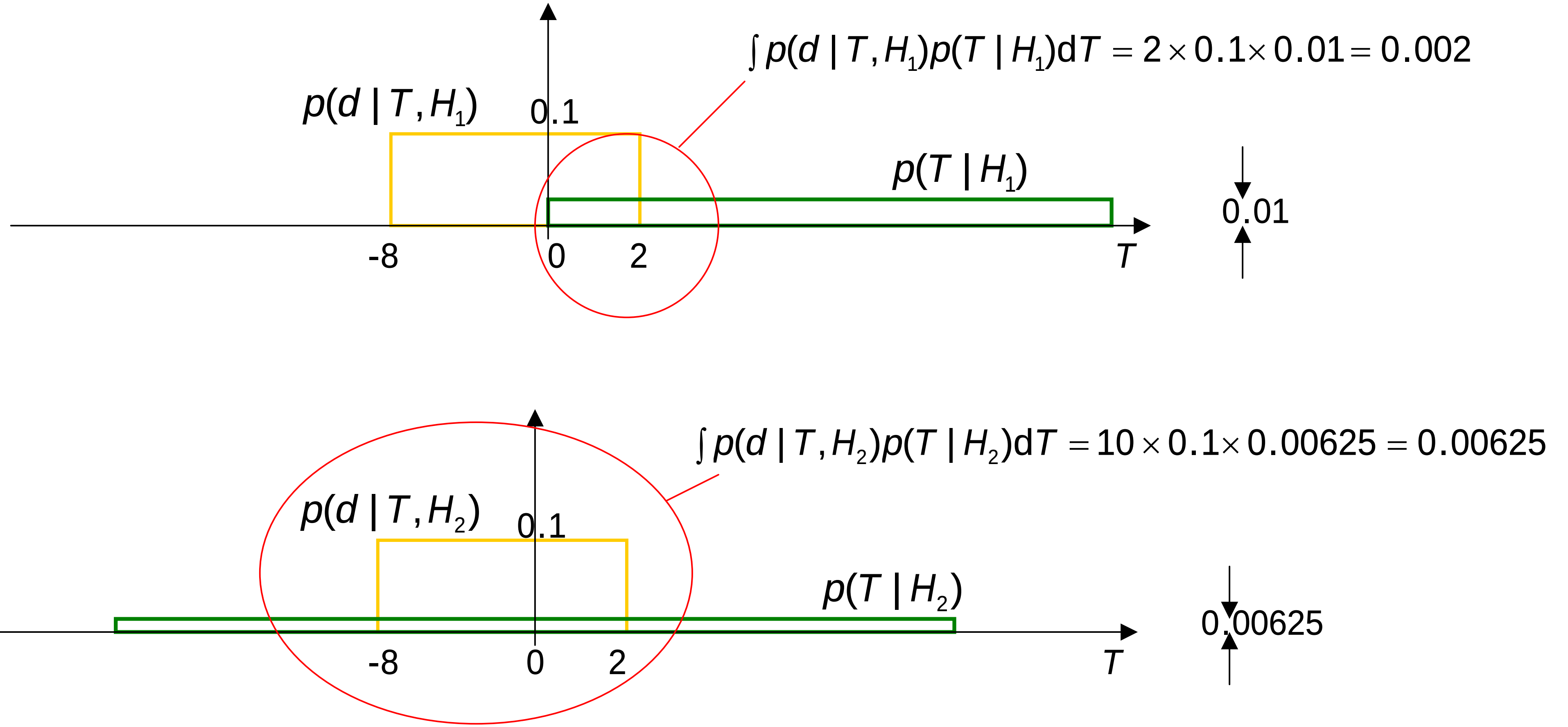
- So under the water hypothesis we have a tighter possible range for the liquid's temperature, but it *may not be water.* In fact, the odds of it being water rather than ethanol are

$$O_{12} = \frac{\text{prob}(H_1 \mid d, I)}{\text{prob}(H_2 \mid d, I)} = \underbrace{\frac{\text{prob}(H_1 \mid I)}{\text{prob}(H_2 \mid I)}}_{\text{prior odds}} \times \underbrace{\frac{\text{prob}(d \mid H_1, I)}{\text{prob}(d \mid H_2, I)}}_{\text{Bayes' factor}} = 1 \times \frac{0.002}{0.00625} = 0.32$$

which means 3:1 in favour of ethanol. Of course this depends on our prior odds too, which we have set to 1. If the choice was between water and *whisky* under the surface of the moon the result would be very different, though the Bayes' factor would be roughly the same!

- Why do we prefer the ethanol option? Because too much of the prior for temperature, assuming water, falls at values that are excluded by the data. In other words, the water hypothesis is *unnecessarily complicated* . Bayes' factors naturally implement Occam's razor in a quantitative way.

Bayesian hypothesis testing



Overlap integral (=evidence) is greater for H_2 (ethanol) than for H_1 (water)