8.2: **Bayesian Statistics

Now that we've seen how easy it is to compute confidence intervals, let's give it a proper probabilistic meaning. To extend probability from the frequentist definition to the Bayesian definition, we need *Bayes' rule*. Bayes' rule is, for events A and B :

$$P(A \mid B)P(B) = P(B \mid A)P(A). \tag{8.2.1}$$

Study Figure 8.4 to convince yourself that Bayes' rule is true. Notice that

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{8.2.2}$$

and

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}. \tag{8.2.3}$$

So, equating $P(A \cap B)$ from each of those two perspectives, we get Bayes' rule.

If we let $A = H$ (hypothesis) and $B = D$ (data), Bayes' rule gives us a way to define the probability of hypothesis through

$$P(H \mid D) = P(D \mid H)\left[\frac{P(H)}{P(D)}\right].$$

The quantity $[P(H)/P(D)]$ is known as the *prior probability* of the data relative to the hypothesis and is something that can be computed in theory if probabilities are assigned in a reasonable manner. The specification of prior probabilities is a contentious issue with the Bayesian approach. Really, it represents a prior *belief*. The quantity $P(D \mid H)$ is what sampling theory, like the central limit theorem, gives and is known as the *likelihood*. Finally the quantity $P(H \mid D)$ is known as the *posterior probability*. Equation (8.2) is an expression about probability distributions as well as individual probabilities (just allow H and D to vary).

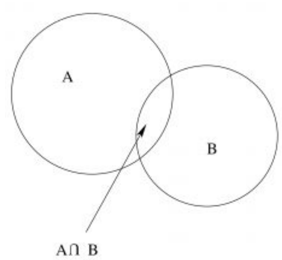

Figure 8.4 : Venn diagram illustration of Bayes rule.

If we assign $[P(H)/P(D)] = 1$ for the prior probability then $P(H \mid D) = P(D \mid H)$. We can switch the roles of D and H! Of course $[P(H)/P(D)] = 1$ is not a probability distribution because the area under a function whose value is always 1 is infinite. The area under a probability distribution must be 1. So $[P(H)/P(D)] = 1$ is an *improper distribution* (as a function of either H or D). But note that an improper distribution times a proper distribution here gives rise to a proper distribution. With this slight of hand, we can give confidence intervals a probabilistic interpretation.