Answers to reading check #11

1. What three quantities must you estimate in order to make use of the theorem that is the subject of this chapter? (You can express them in general terms, or in the specific terms of finding panties in a man's underwear drawer.)

   To use Bayes theorem, you must know: 1. The probability of the panties being present, assuming your partner is *indeed* cheating. 2. The probability of your partner cheating, for any reason. 3. The overall probability of the panties being present for any reason.

   More generally, you must know: 1. The likelihood. 2. The prior. 3. The evidence. These three give you the posterior probability, which is what you want:

   $$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

   $$P(\text{cheating}|\text{panties}) = \frac{P(\text{panties}|\text{cheating}) \times P(\text{cheating})}{P(\text{panties})}$$

2. Professional sports gambler Haralabos Voulgaris big secret is that he doesn't have a big secret. What does he have instead?

   A lot of little secrets. (He's a fox, not a hedgehog.)

3. What's the fundamental difference between Bayesianism vs. Frequentism?

   Bayesian reasoning conceives of probabilities as subjective degrees of belief, whereas Frequentism insists that they can only come from the long-range prevalence of actual observations. Frequentism also asserts that the uncertainty in statistics is solely due to sampling error, whereas Bayesianism admits that there are many biases present in any judgment about probabilities.

4. On what scientific question did R. A. Fisher make a grievous error of judgment in the last years of his life, which helped demonstrate the limitations of his statistical approach?

   Whether smoking causes lung cancer.

5. John P. A. Ioannidis observed that we have seen explosive growth in the availability of information in the last few decades, and we can now measure millions and millions of potentially interesting variables. What disadvantage to this does he mention?

   Our predictions are probably *more* prone to failure in the era of Big Data. As there is an increase in the amount of available information, there's also an exponential increase in the number of hypotheses to investigate. But the number of *meaningful* relationships in the data is orders of magnitude smaller: there isn't any more truth in the world now than there was before the Internet (or the printing press).