# Statistics 101
# Data Analysis and Statistical Inference

## Answers to Extra Problems on Bayesian Stats

---

### 1. The weather, the weather

Let R = it rains.
Let C = clouds roll in.
We want Pr (R | C)

From the problem, we know that
Pr (R) = 0.30.
Pr(C | R) = 0.95
Pr(C | not R) = 0.25.

Hence, using Bayes rule, we have:

Pr (R | C) = Pr (R and C) / Pr (C) = Pr (C | R) Pr (R) / Pr (C) = (.95)(.30) / Pr(C).

Now, Pr (C) = Pr ( C and R) + Pr (C and not R) = (.95)(.30) + Pr (C | not R) Pr (not R) = (.95)(.30) + (.25)(.70) = 0.46

Hence, Pr (R | C ) = (.95)(.30) / .46 = .619.

There is a 61.9% chance that it will rain, given that clouds rolled in during the morning.

### 2. Auditing tax returns

Let E = the return has an error.
Let F = return is flagged by computer.
We want Pr (E | F)

From the problem, we know that
Pr (E) = 0.15.
Pr(F | E) = 0.80
Pr(F | not E) = 0.05.

Hence, using Bayes rule, we have:

Pr (E | F) = Pr (E and F) / Pr (F) = Pr (F | E) Pr (E) / Pr (F) = (.80)(.15) / Pr(F).

Now, Pr (F) = Pr ( F and E) + Pr (F and not E) = (.80)(.15) + Pr (F | not E) Pr (not E) = (.80)(.15) + (.05)(.85) = 0.1625

Hence, Pr (E | F ) = (.80)(.15) / .1625 = .7385.

There is a 73.85% chance that the return has an error, given that the computer flagged it. Notice that this is a big increase from 15%. The computer really helps identify erroneous returns.

## 3. Paternity suits

Let A = alleged father is the real father.
Let B = child has type B blood.
We want Pr (A | B)

From the problem, we know that
Pr (A) = 0.75.
Pr(B | A) = 0.50
Pr(B | not A) = 0.09.

Hence, using Bayes rule, we have:

Pr (A | B) = Pr (A and B) / Pr (B) = Pr (B | A) Pr (A) / Pr (B) = (.50)(.75) / Pr(B).

Now, Pr (B) = Pr ( B and A) + Pr (B and not A) = (.50)(.75) + Pr (B | not A) Pr (not A) = (.50)(.75) + (.09)(.25) = 0.3975

Hence, Pr (A | B ) = (.50)(.75) / .3975 = .9434.

There is a 94.34% chance that the alleged father is the real father, given the child is blood type B.

## 4. Differences between Bayesian and classical inference

a) In classical inference, the probability, Pr(mu > 1400), is a number strictly bigger than zero and strictly less than one.

False. In classical inference, mu is not treated as random. Rather, it is some fixed number. Hence, mu is either greater than 1400 or less than 1400. This implies that Pr(mu > 1400) must equal zero or it must equal one. It cannot be a number in between zero and one.

b) In Bayesian inference, the probability, Pr(mu > 1400), is a number strictly bigger than zero and strictly less than one.

True. In Bayesian inference, mu is treated as random. We make probability statements about mu by using its posterior distribution. Hence, Pr(mu > 1400) is some number between zero and one.

c) In classical inference, our best guess at mu is its maximum likelihood estimate.

True. For the normal curve, the maximum likelihood estimate of mu equals the sample mean of the data.

d) If you have very strong prior beliefs about mu, the Bayesian's best guess at mu will be affected by those beliefs.

True. The Bayesian's best guess at mu combines the prior information about mu and the data. For example, for the normal curve, the Bayesian's best guess at mu is a weighted average of the sample mean and the prior mean.

e) If you draw a likelihood function for mu, the best guess at mu is the number corresponding to the top of the hill in the likelihood function.

True. Maximum likelihood estimates are those which maximize the likelihood function, i.e., have the largest values of the likelihood function.

---