Lesson 6

Back to Week 3

6/6 points earned (100%)

Quiz passed!

1 / 1 points

1.

For Questions 1-2, consider the following experiment:

Suppose you are trying to calibrate a thermometer by testing the temperature it reads when water begins to boil. Because of natural variation, you take several measurements (experiments) to estimate θ, the mean temperature reading for this thermometer at the boiling point.

You know that at sea level, water should boil at 100 degrees Celsius, so you use a precise prior with $P(\theta = 100) = 1$. You then observe the following five measurements: 94.6 95.4 96.2 94.9 95.9.

- What will the posterior for θ look like?

○ Most posterior probability will be concentrated near the sample mean of 95.4 degrees Celsius.

○ Most posterior probability will be spread between the sample mean of 95.4 degrees Celsius and the prior mean of 100 degrees Celsius.

○ The posterior will be $\theta = 100$ with probability 1, regardless of the data.

Correct Response
Because all prior probability is on a single point (100 degrees Celsius), the prior completely dominates any data. If we are 100% certain of the outcome before the experiment, we learn nothing by performing it.

Clearly this was a poor choice of prior, especially in light of the data we collected.

○ None of the above.

---

1 / 1
points

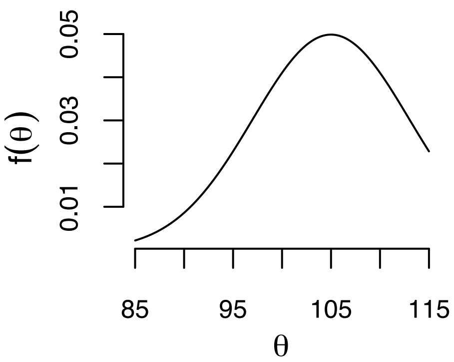
2.

Thermometer:

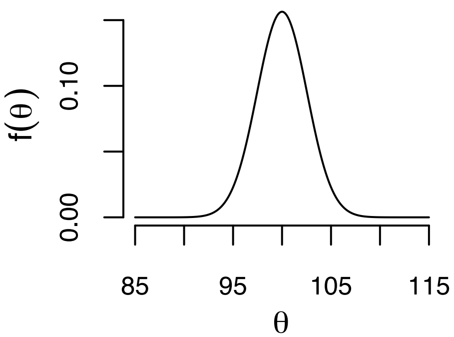
Suppose you believe before the experiments that the thermometer is biased high, so that on average it would read 105 degrees Celsius, and you are 95% confident that the average would be between 100 and 110.

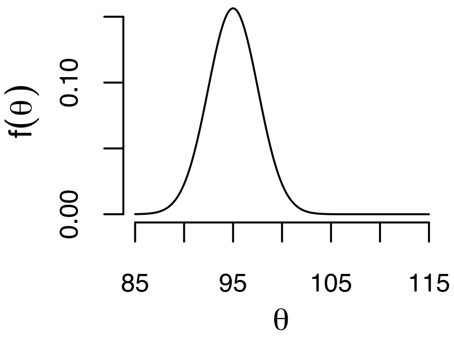
- Which of the following prior PDFs most accurately reflects this prior belief?

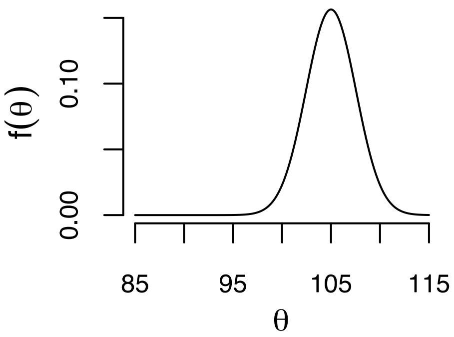
○



○







Correct Response

The prior mean is 105 degrees Celsius and approximately 95% of the prior probability is assigned to the interval (100, 110).

1 / 1
points

Recall that for positive integer n, the gamma function has the following property: $\Gamma(n) = (n-1)!$.

3. What is the value of $\Gamma(6)$?

120

Correct Response
This is $\Gamma(6) = 5! = 120$.

1 / 1
points

4.
Find the value of the normalizing constant, c, which will cause the following integral to evaluate to 1.

$\int_0^1 c \cdot z^3 (1-z)^1 dz.$

Hint: Notice that this is proportional to a beta density. We only need to find the values of the parameters α and β and plug those into the usual normalizing constant for a beta density.

$\frac{\Gamma(4+2)}{\Gamma(4)\Gamma(2)} = \frac{5!}{3!1!} = 20$

Correct Response
$\alpha = 3 + 1$ and $\beta = 1 + 1$.

- $\frac{\Gamma(1)}{\Gamma(z)\Gamma(1-z)} = \frac{0!}{(z-1)!1!}$
- $\frac{\Gamma(3+1)}{\Gamma(3)\Gamma(1)} = \frac{3!}{2!0!} = 3$

---

1 / 1 points

5.

Consider the coin-flipping example from Lesson 5. The likelihood for each coin flip was Bernoulli with probability of heads θ, or $f(y \mid \theta) = \theta^y (1-\theta)^{1-y}$ for $y = 0$ or $y = 1$, and we used a uniform prior on θ.

Recall that if we had observed $Y_1 = 0$ instead of $Y_1 = 1$, the posterior distribution for θ would have been $f(\theta \mid Y_1 = 0) = 2(1-\theta) I_{\{0 \le \theta \le 1\}}$. Which of the following is the correct expression for the posterior predictive distribution for the next flip $Y_2 \mid Y_1 = 0$?

- $f(y_2 \mid Y_1 = 0) = \int_0^1 2(1-\theta) d\theta$ for $y_2 = 0$ or $y_2 = 1$.
- $f(y_2 \mid Y_1 = 0) = \int_0^1 \theta^{y_2} (1-\theta)^{1-y_2} d\theta$ for $y_2 = 0$ or $y_2 = 1$.
- $f(y_2 \mid Y_1 = 0) = \int_0^1 2\theta^{y_2} (1-\theta)^{1-y_2} d\theta$ for $y_2 = 0$ or $y_2 = 1$.
- $f(y_2 \mid Y_1 = 0) = \int_0^1 \theta^{y_2} (1-\theta)^{1-y_2} 2(1-\theta) d\theta$ for $y_2 = 0$ or $y_2 = 1$.

Correct Response

This is just the integral over likelihood $\times$ posterior. This expression simplifies to

$$\int_0^1 2\theta^{y_2} (1-\theta)^{2-y_2} d\theta I_{\{y_2 \in \{0,1\}\}} = \frac{2}{\Gamma(4)} \Gamma(y_2+1)\Gamma(3-y_2) I_{\{y_2 \in \{0,1\}\}}$$

$$= \frac{2}{3} I_{\{y_2 = 0\}} + \frac{1}{3} I_{\{y_2 = 1\}}$$

1 / 1
points

6.
The prior predictive distribution for X when θ is continuous is given by $\int f(x \mid \theta) \cdot f(\theta) d\theta$. The analogous expression when θ is discrete is $\sum_\theta f(x \mid \theta) \cdot f(\theta)$, adding over all possible values of θ.

Let's return to the example of your brother's loaded coin from Lesson 5. Recall that he has a fair coin where heads comes up on average 50% of the time ($p = 0.5$) and a loaded coin ($p = 0.7$). If we flip the coin five times, the likelihood is binomial: $f(x \mid p) = \binom{5}{x} p^x (1-p)^{5-x}$ where X counts the number of heads.

Suppose you are confident, but not sure that he has brought you the loaded coin, so that your prior is $f(p) = 0.9 I_{\{p=0.7\}} + 0.1 I_{\{p=0.5\}}$. Which of the following expressions gives the prior predictive distribution of X?

- $f(x) = \binom{5}{x} .7^x (.3)^{5-x} (.5) + \binom{5}{x} .5^x (.5)^{5-x} (.5)$
- $f(x) = \binom{5}{x} .7^x (.3)^{5-x} (.1) + \binom{5}{x} .5^x (.5)^{5-x} (.9)$
- $f(x) = \binom{5}{x} .7^x (.3)^{5-x} (.9) + \binom{5}{x} .5^x (.5)^{5-x} (.1)$

Correct Response
This is a weighted average of binomials, with weights being your prior probabilities for each scenario (loaded or fair).

- $f(x) = \binom{5}{x} .7^x (.3)^{5-x} + \binom{5}{x} .5^x (.5)^{5-x}$