Lesson 10

LATEST SUBMISSION GRADE

100%

1. **For Questions 1-6, consider the thermometer calibration problem from the quiz in Lesson 6.**

1 / 1 point

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 n independent measurements (experiments) to estimate θ, the mean temperature reading for this thermometer at the boiling point. Assume a normal likelihood for these data, with mean θ and known variance $\sigma^2 = 0.25$ (which corresponds to a standard deviation of 0.5 degrees Celsius).

Suppose your prior for θ is (conveniently) the conjugate normal. You know that at sea level, water should boil at 100 degrees Celsius, so you set the prior mean at $m_0 = 100$.

- If you specify a prior variance s_0^2 for θ, which of the following accurately describes the model for your measurements Y_i, $i = 1, \ldots, n$?

- (selected) $Y_i \mid \theta \overset{\text{iid}}{\sim} \text{N}(\theta, 0.25)\,;\, \theta \sim \text{N}(100, s_0^2)$
- $Y_i \mid \sigma^2 \overset{\text{iid}}{\sim} \text{N}(100, \sigma^2)\,;\, \sigma^2 \sim \text{Inverse-Gamma}(0.25, s_0^2)$
- $Y_i \mid \theta, \sigma^2 \overset{\text{iid}}{\sim} \text{N}(\theta, \sigma^2)\,;\, \sigma^2 \sim \text{Inverse-Gamma}(100, s_0^2)$
- $Y_i \mid \theta \overset{\text{iid}}{\sim} \text{N}(100, 0.25)\,;\, \theta \sim \text{N}(\theta, s_0^2)$
- $Y_i \mid \theta \overset{\text{iid}}{\sim} \text{N}(\theta, 100)\,;\, \theta \sim \text{N}(0.25, s_0^2)$

Correct
This is a normal likelihood with known variance, and normal prior on the mean.

2. Thermometer calibration:

1 / 1 point

You decide you want the prior to be equivalent (in effective sample size) to one measurement.

- What value should you select for s_0^2 the prior variance of θ? Round your answer to two decimal places.

0.25

Correct
The prior effective sample size is $\frac{\sigma^2}{s_0^2} = \frac{0.25}{0.25} = 1$.

3. Thermometer calibration:

1 / 1 point

You collect the following $n = 5$ measurements: (94.6, 95.4, 96.2, 94.9, 95.9).

- What is the posterior distribution for θ?

- $\text{N}(95.41, 0.250)$
- $\text{N}(95.41, 0.042)$
- $\text{N}(100, 0.250)$
- (selected) $\text{N}(96.17, 0.042)$
- $\text{N}(95.41, 24)$
- $\text{N}(96.17, 24)$

Correct
Plugging all relevant quantities (including $\bar{y} = 95.4$) into the update formula in Lesson 10.1, the posterior mean is $\frac{2308}{24}$ and the posterior variance is $\frac{1}{24}$.

4. Thermometer calibration:

1 / 1 point

- Use R or Excel to find the upper end of a 95% equal-tailed credible interval for θ.

96.57167

Correct

This is the 0.975 quantile of the posterior distribution.

In R:

```
qnorm(p=0.975, mean=96.17, sd=sqrt(0.042))
```

In Excel:

```
= NORM.INV(0.975, 96.17, SQRT(0.042))
```

where probability=0.975, mean=96.17, and standard_dev=SQRT(0.042).

5. Thermometer calibration:

1 / 1 point

- After collecting these data, is it reasonable to conclude that the thermometer is biased toward low values?

- Yes, we have $P(\theta < 100 \mid \mathbf{y}) > 0.9999$.
- Yes, we have $P(\theta > 100 \mid \mathbf{y}) > 0.9999$.
- No, we have $P(\theta < 100 \mid \mathbf{y}) < 0.0001$.
- No, we have $P(\theta = 100 \mid \mathbf{y}) = 0$.

Correct

Using the posterior distribution $\mathrm{N}(a, b)$,

In R:

```
pnorm(q=100, mean=a, sd=sqrt(b))
```

In Excel:

```
= NORM.DIST(100, a, SQRT(b), TRUE)
```

where x=100, mean=a, standard_dev=SQRT(b), and cumulative=TRUE.

6. Thermometer calibration:

1 / 1 point

- What is the posterior predictive distribution of a single future observation Y^*?

- $\mathrm{N}(96.17, 0.292)$
- $\mathrm{N}(95.41, 0.50)$
- $\mathrm{N}(96.17, 0.042)$
- $\mathrm{N}(100, 0.50)$
- $\mathrm{N}(95.41, 0.292)$

Correct

This is the posterior distribution with the variance increased by the value of known data variance.

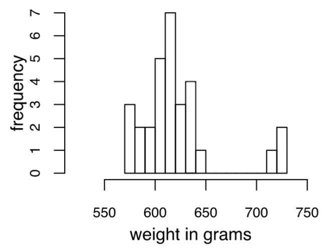
7. **For Questions 7-10, consider the following scenario:**

1 / 1 point

Your friend moves from city A to city B and is delighted to find her favorite restaurant chain at her new location. After several meals, however, she suspects that the restaurant in city B is less generous. She decides to investigate.

She orders the main dish on 30 randomly selected days throughout the year and records each meal's weight in grams. You still live in city A, so you assist by performing the same experiment at your restaurant. Assume that the dishes are served on identical plates (measurements subtract the plate's weight), and that your scale and your friend's scale are consistent.

The following histogram shows the 30 measurements from Restaurant B taken by your friend.



- Is it reasonable to assume that these data are normally distributed?

- ( ) Yes, the distribution appears to follow a bell-shaped curve.
- ( ) Yes, the data are tightly clustered around a single number.
- ( ) No, the first bar to the left of the peak is not equal in height to he first bar to the right of the peak.
- (•) No, there appear to be a few extreme observations (outliers).

Correct

The three points above 700 are about five (sample) standard deviations above the (sample) mean.

8. Restaurants: **1 / 1 point**

Your friend investigates the three observations above 700 grams and discovers that she had ordered the incorrect meal on those dates. She removes these observations from the data set and proceeds with the analysis using $n = 27$.

She assumes a normal likelihood for the data with unknown mean μ and unknown variance σ^2. She uses the model presented in Lesson 10.2 where, conditional on σ^2, the prior for μ is normal with mean m and variance σ^2/w. Next, the marginal prior for σ^2 is $\text{Inverse-Gamma}(a, b)$.

Your friend's prior guess on the mean dish weight is 500 grams, so we set $m = 500$. She is not very confident with this guess, so we set the prior effective sample size $w = 0.1$. Finally, she sets $a = 3$ and $b = 200$.

We can learn more about this inverse-gamma prior by simulating draws from it. If a random variable X follows a $\text{Gamma}(a, b)$ distribution, then $\frac{1}{X}$ follows an $\text{Inverse-Gamma}(a, b)$ distribution. Hence, we can simulate draws from a gamma distribution and take their reciprocals, which will be draws from an inverse-gamma.

To simulate 1000 draws in R (replace a and b with their actual values):

```
z <- rgamma(n=1000, shape=a, rate=b)
x <- 1/z
```

To simulate one draw in Excel (replace a and b with their actual values):

```
= 1 / GAMMA.INV( RAND(), a, 1/b )
```

where probability=RAND(), alpha=a, and beta=1/b. Then copy this formula to obtain multiple draws.

- Simulate a large number of draws (at least 300) from the prior for σ^2 and report your approximate prior mean from these draws. It does not need to be exact.

98.5984

Correct

The actual prior mean for σ^2 is $\frac{b}{a-1} = \frac{200}{2} = 100$. The prior variance for σ^2 is $\frac{b^2}{(a-1)^2(a-2)} = 10,000$.

9. Restaurants: **1 / 1 point**

With the $n = 27$ data points, your friend calculates the sample mean $\bar{y} = 609.7$ and sample variance $s^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2 = 401.8$.

Using the update formulas from Lesson 10.2, she calculates the following posterior distributions:

$\sigma^2 \mid \mathbf{y} \sim \text{Inverse-Gamma}(a', b')$

$\mu \mid \sigma^2, \mathbf{y} \sim \text{N}(m', \frac{\sigma^2}{w+n})$

where

$a' = a + \frac{n}{2} = 3 + \frac{27}{2} = 16.5$

$b' = b + \frac{n-1}{2} s^2 + \frac{wn}{2(w+n)} (\bar{y} - m)^2 = 200 + \frac{27-1}{2} 401.8 + \frac{0.1 \cdot 27}{2(0.1+27)} (609.7 - 500)^2 = 6022.9$

$m' = \frac{n\bar{y} + wm}{w+n} = \frac{27 \cdot 609.7 + 0.1 \cdot 500}{0.1+27} = 609.3$

$w = 0.1$, and $w + n = 27.1$.

To simulate draws from this posterior, begin by drawing values for σ^2 from its posterior using the method from the preceding question. Then, plug these values for σ^2 into the posterior for μ and draw from that normal distribution.

To simulate 1000 draws in R:

```
z <- rgamma(1000, shape=16.5, rate=6022.9)
sig2 <- 1/z
mu <- rnorm(1000, mean=609.3, sd=sqrt(sig2/27.1))
```

To simulate one draw in Excel:

```
= 1 / GAMMA.INV( RAND(), 16.5, 1/6022.9 )
```

gets saved into cell A1 (for example) as the draw for σ^2. Then draw

```
= NORM.INV( RAND(), 609.3, SQRT(A1/27.1) )
```

where probability=RAND(), mean=609.3, standard_dev=SQRT(A1/27.1), and A1 is the reference to the cell containing the draw for σ^2. Then copy these formulas to obtain multiple draws.

We can use these simulated draws to help us approximate inferences for μ and σ^2. For example, we can obtain a 95% equal-tailed credible for μ by calculating the quantiles/percentiles of the simulated values.

In R:

```
quantile(x=mu, probs=c(0.025, 0.975))
```

In Excel:

```
= PERCENTILE.INC( A1:A500, 0.025 )
= PERCENTILE.INC( A1:A500, 0.975 )
```

where array=A1:A500 (or the cells where you have stored samples of μ) and k=0.025 or 0.975.

- Perform the posterior simulation described above and compute your approximate 95% equal-tailed credible interval for μ. Based on your simulation, which of the following appears to be the actual interval?

- (602, 617)
- (608, 610)
- (582, 637)
- (245, 619)

Correct

This is the actual interval, calculated from the exact marginal posterior (t distribution) for μ.

10. Restaurants: 1 / 1 point

You complete your experiment at Restaurant A with $n = 30$ data points, which appear to be normally distributed. You calculate the sample mean $\bar{y} = 622.8$ and sample variance $s^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2 = 403.1$.

Repeat the analysis from Question 9 using the same priors and draw samples from the posterior distribution of σ_A^2 and μ_A (where the A denotes that these parameters are for Restaurant A).

Treating the data from Restaurant A as independent from Restaurant B, we can now attempt to answer your friend's original question: is restaurant A more generous? To do so, we can compute posterior probabilities of hypotheses like $\mu_A > \mu_B$. This is a simple task if we have simulated draws for μ_A and μ_B. For $i = 1, \ldots, N$ (the number of simulations drawn for each parameter), make the comparison $\mu_A > \mu_B$ using the ith draw for μ_A and μ_B. Then count how many of these return a TRUE value and divide by N, the total number of simulations.

In R (using 1000 simulated values):

```
sum( muA > muB ) / 1000
```

or

```
mean( muA > muB )
```

In Excel (for one value):

```
= IF(A1 > B1, 1, 0)

```

where the first argument is the logical test which compares the value of cell A1 with that of B1, 1=value_if_true, and 0=value_if_false. Copy this formula to compare all μ_A, μ_B pairs. This will yield a column of binary (0 or 1) values, which you can sum or average to approximate the posterior probability.

- Would you conclude that the main dish from restaurant A weighs more than the main dish from restaurant B on average?

- (x) Yes, the posterior probability that $\mu_A > \mu_B$ is at least 0.95.
- ( ) Yes, the posterior probability that $\mu_A > \mu_B$ is less than 0.05.
- ( ) No, the posterior probability that $\mu_A > \mu_B$ is at least 0.95.
- ( ) No, the posterior probability that $\mu_A > \mu_B$ is less than 0.05.

Correct

This is fairly strong evidence that the mean weight of the dish from Restaurant A is greater than the mean weight of the dish from Restaurant B.