1. Fill in the blank: Under a **0/1** loss function, the summary statistic that minimizes the posterior expected loss is the ____________________ of the posterior.

1 / 1 point

- ( ) Mean
- ( ) Median
- (•) Mode

Correct

Correct answer. The mode is the summary statistic that minimizes the posterior expected loss under the 0/1 loss function.

This question refers to the following learning objective(s):

- Understand the concept of loss functions and how they relate to Bayesian decision making.

2. You are employed in the Human Resources department of a company and are asked to predict the number of employees that will quit their jobs in the coming year. Based on a Bayesian analysis of historical data, you determine the posterior predictive distribution of the number of quitters to follow a Poisson($\lambda = 10$) distribution. Given a quadratic loss function, what is the prediction that minimizes posterior expected loss?

1 / 1 point

- ( ) 9
- (•) 10
- ( ) 11

Correct

10 is the posterior mean of Poisson($\lambda = 10$). Since the loss function is quadratic, the mean of the posterior distribution minimizes the posterior expected loss.

This question refers to the following learning objective(s):

- Make optimal decisions given a posterior distribution and a loss function.

3.

1 / 1 point

True or False: If the posterior distribution is a Binomial distribution, the estimate that minimizes posterior expected loss is the same, regardless of whether the loss function is 0/1, linear, or quadratic.

- (●) False
- ( ) True

Correct

Correct answer. For a Binomial distribution with n trials and success rate p, the mean is np. The mean, median, and mode will only be the same when np is an integer, since median and mode are always whole numbers for Binomial distributions.

This question refers to the following learning objective(s):

- Make optimal decisions given a posterior distribution and a loss function.

4. Suppose that you are trying to decide whether a coin is biased towards heads $(p = 0.75)$ or tails $(p = 0.25)$. If you decide the coin is biased towards heads but it is not, you incur a loss of 10. On the other hand, if you decide the coin is biased towards tails but it is not, you incur a loss of 100. If you make the correct choice, you will not incur any losses. At what posterior probability of the coin being **biased towards heads** will you be indifferent between the two decisions? **1 / 1 point**

- ( ) 0.0775
- (●) 0.091
- ( ) 0.5
- ( ) 0.325

Correct

Correct answer. Let $P(p = 0.75)$ be the posterior probability of the coin being head-biased. The expected loss when you decide the coin is head-biased when it is not is $10 \times (1 - P(p = 0.75))$. The expected loss when you decide the coin is tail-biased when it is not is $100 \times P(p = 0.75)$. Equate the two expected losses and solve for $P(p = 0.75)$.

This question refers to the following learning objective(s):

- Decide between hypotheses given a loss function.

5. Which of the following statements is false?

1 / 1 point

- (selected) A Bayes factor of less than .01 suggests that the evidence in favor of one of the hypotheses is barely worth mentioning
- We can update prior odds to posterior odds using Bayes factors
- A Bayes factor of greater than 100 suggests strong evidence in favor of one of the hypotheses
- The Bayes factor represents the ratio of the marginal likelihoods of observing the data under the two hypotheses

✓ **Correct**

A Bayes factor of less than .01 will yield strong evidence in favor of one of the hypotheses.

6. Suppose your posterior distribution for the population mean μ is a Student t distribution centered at 10 and you find a 95% Highest Posterior Density interval with (L, U). **Which of the following is true:**

1 / 1 point

- L is two standard deviations below 10 while U is two standard deviations above 10
- (selected) The midpoint between L and U is 10
- The $P(\mu \geq U \mid \text{data}) = .05$
- The $P(\mu \leq L \mid \text{data}) = .05$

Correct

When a distribution is symmetric and unimodal the highest probability will happen at the center of the distribution leading to an interval that is symmetric about the center

7. When modeling data $Y_i \sim N(\mu, \sigma^2)$, when both μ and σ^2 are unknown, it is often useful to use a Normal-Gamma prior distribution. Our joint prior distribution $p(\mu, \sigma^2)$ can be written as: **1 / 1 point**

$$\mu \mid \sigma^2 \ \sim \ N(m_0, \sigma^2/n_0)$$

$$1/\sigma^2 \ \sim \ \text{Gamma}(v_0/2, v_0 s_0^2/2),$$

Which of the following hyperparameters can be described as our initial guess about the variance parameter σ^2

- ( ) m_0.
- ( ) n_0.
- ( ) v_0.
- (•) s_0^2

Correct

s_0^2 represents our initial guess about the variance parameter σ^2, while v_0 represents our prior degrees of freedom (which can be thought of as a measure of confidence in our guess).

This question refers to the following learning objective(s):

- Understanding of the meaning of hyperparameters of Normal-Gamma conjugate families.

8. True or False: If data come from a Normal distribution where both the mean and variance are unknown and we use a Jeffreys-Zellner-Siow prior, we need to use simulation techniques to create credible intervals for μ since there is no closed form representation of the posterior distribution for μ. **1 / 1 point**

- (•) True

- [ ] False

Correct

The Jeffreys-Zellner-Siow prior is a mixture of conjugate priors and the prior has a closed form representation, however the posterior distribution does not. To create credible intervals we use simulation based methods.

9. In the *Playing Computer Game During Lunch Affects Fullness, Memory for Lunch, and Later Snack Intake* study, researchers evaluated the relationship between being distracted and recall of food consumed and snacking. The sample of this study consisted of 44 volunteer patients, randomized into 2 groups with equal size. One group was asked to play solitaire on the computer while eating and was asked to win as many games as possible, and the other group was asked to eat without any distractions, focusing on what they're eating and thinking about the taste of the food. Both groups were provided the same amount of lunch and after lunch, they were offered cookies to snack on.

1 / 1 point

The distracted group snacked an average of $\bar{Y} = 52.1$ grams of cookies, with sample standard deviation $s = 45.1$ grams and sample size $n = 22$.

Under the unknown mean and variance, the researcher determined that the 95% highest probability density credible interval for μ was $[32.1, 72.1]$ grams and that the posterior mean (median) was 52.1 grams.

Based on the above, which statement is true:

- [ ] There is a 50% chance that the distracted eaters will consume 52.1 grams of cookies.
- [ ] There is a 95% chance a distracted eater will consume between 32.1 and 72.1 grams of cookies
- [ ] The 95% credible interval means 95% of random samples of 22 distracted eaters will yield intervals that contain the true mean of average snack intake level.
- [x] There is a 95% chance that on average, distracted eaters consume between 32.1 and 72.1 grams of cookies
- [ ] The 95% Bayesian credible interval is different from the 95% frequentist confidence interval for μ in this case.

Correct

μ describes what happens on average in the population; given the data we expect that there is 95% chance that the average consumption is in this interval

10. In the *Playing Computer Game During Lunch Affects Fullness, Memory for Lunch, and Later Snack Intake* study, researchers evaluated the relationship between being distracted and recall of food consumed and snacking. The sample of this study consisted of 44 volunteer patients, randomized into 2 groups with equal size. One group was asked to play solitaire on the computer while eating and was asked to win as many games as possible, and the other group was asked to eat without any distractions, focusing on what they're eating and thinking about the taste of the food. Both groups were provided the same amount of lunch and after lunch, they were offered cookies to snack on. Consumption of cookies (in grams) after lunch was measured for each of the volunteers. In the treatment group, mean consumption was 52.1 grams with standard deviation 45.1 grams. In the control group, mean consumption was 27.1 grams with standard deviation 26.4 grams. The research was interested in testing the hypothesis that the average consumption was different in the two groups.

0 / 1 point

Based on the information provided, which of the following assumptions is needed for testing the hypothesis that the average consumption is the same in the two groups versus that the average consumption is not the same?

- ○ Individuals within each group d cannot influence the cookie consumption of others in the same group, i.e. consumption is independent within groups
- ○ Cookie consumption is independent between groups
- ○ Cookie consumption in grams is normally distributed
- ○ The variability of consumption is the same in both groups
- ○ All of the above
- ◉ We do not need any assumptions for Bayesian hypothesis testing, so none of the above

✗ **Incorrect**

Review the video **Comparing Two Paired Means Using Bayes Factors**

11.

1 / 1 point

1000 students were randomly sampled from the High School and Beyond survey. Each student took a reading and a writing test leading to paired observations. We want to test the hypothesis whether the average score of the reading test, μ_R, is the same as the average score of the writing test, μ_W. Since a student's reading score is not likely to be independent of his/her own writing score, we test hypotheses about the two scores, $\mu = \mu_W - \mu_R$,

$H_1 : \mu = 0$

$H_2 : \mu \neq 0$

The resulting Bayes factor for comparing H_1 to H_2 was 3.505.

We can conclude using the Jeffrey's scale of evidence that:

- ○ The Bayes factor provides positive evidence against the hypothesis that the mean reading and mean writing test scores are the same
- ○ The Bayes factor provides strong evidence against H_1
- ◉ The Bayes factor provides positive evidence against the mean of the writing scores being different from the mean reading score.The Bayes factor is 3.505, which is larger than 3. So it provides positive evidence against H_2 according to Jeffrey's scale of evidence.
- ○ The Bayes factor provides strong evidence against H_2

✓ **Correct**

12. For Hypothesis tests about a Normal mean μ or the difference in two means where $\mu = \mu_1 - \mu_2$, where $H_1 : \mu = 0$ versus $H_2 : \mu \neq 0$ which prior distribution is recommended for μ under H_2 for most circumstances? **1 / 1 point**

- ○ A non-informative Normal distribution for μ with a really large variance, σ^2/n_0 by taking n_0 to be very small.
- ○ The reference prior for μ (a uniform or flat distribution)
- ◉ A Cauchy distribution for μ
- ○ All of the above

Correct

The Cauchy prior has heavy tails that accommodates situations when the prior mean is far from the actual sample mean. See the paradox discussion in the **Comparing Two Paired Means Using Bayes' Factors** video.