MIDTERM EXAM - Solutions

STAD91 WINTER 2026
University of Toronto Scarborough

Name:
Student #:

Exam duration: 120 minutes
No calculators will be allowed during the midterm exam.

Please check that your exam has 13 pages, including this one. The total possible number of points is 100.

Read the following instructions carefully:

1. Exam is closed book and internet. You can use an optional handwritten aid sheet - A4 (or 8.5" $\times$ 11") double-sided.
2. If a question asks you to do some calculations, you must **show your work** for full credit.
3. Conceptual questions do not require long answers.
4. You will write your answers to each question in the space provided on the exam sheet. If you require additional paper, simply raise your hand.
5. After solving each question, you should write your answers immediately. Do not wait last minute to write them all at once.
6. Do not share the exam with anyone or in any platform!
7. Lastly, enjoy the problems!!!

1. Conceptual Questions (16 points)

Answer by **True** or **False**. No justification needed.

1. (3 points) If the prior distribution $\pi(\theta)$ is improper, then the posterior distribution $\pi(\theta \mid \mathbf{X})$ is necessarily improper. **Correction: False.**

2. (3 points) The Bayes estimator associated with the absolute error loss function $\ell(\theta, T) = |\theta - T|$ is the posterior mean. **Correction: False.**

3. (3 points) In a hypothesis test H_0 vs H_1, if the Bayes Factor $B^{\pi}_{0/1} = 10$, this provides strong evidence in favor of H_1. **Correction: False.** The Bayes Factor $B^{\pi}_{0/1}$ is the ratio of the marginal likelihood of H_0 to that of H_1. A value of $B^{\pi}_{0/1} = 10$ means the data is 10 times more likely under H_0 than under H_1, which provides evidence **in favor of** H_0, not H_1.

4. (3 points) If a Bayes estimator is unique and has constant risk (i.e., $R(\theta, T) = c$ for all θ), then it is minimax. **Correction: True.**

5. (4 points) In a linear regression model with Gaussian noise, obtaining the Maximum A Posteriori (MAP) estimator using a centered Gaussian prior on the coefficients is equivalent to performing Ridge regression. **Correction: True.**

2. Bayes and Minimax Risk (45 points)

For $a > 0$ and $b > 0$, the Inverse-Beta distribution, denoted $IB(a, b)$, is the distribution whose density is given by

$$\theta \mapsto \frac{1}{B(a,b)}\theta^{-a-b}(\theta-1)^{b-1}\mathbb{1}_{(1,+\infty)}(\theta),$$

where $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and Γ is the gamma function, satisfying $\Gamma(x+1) = x\Gamma(x)$ for all $x > 0$. The distribution $IB(a, b)$ is the distribution of the inverse of a random variable following a $\text{Beta}(a, b)$ distribution. Furthermore, if $Y \sim IB(a, b)$ with $a > 2$, we have

$$\mathbb{E}[Y] = \frac{a+b-1}{a-1} \quad \text{and} \quad \text{Var}(Y) = \frac{(a+b-1)b}{(a-1)^2(a-2)}.$$

We consider the model:

$$\boldsymbol{\theta} \sim \Pi := IB(a, 2)$$
$$\mathbf{X} = (X_1, \ldots, X_n) \mid \boldsymbol{\theta} = \theta \sim \mathcal{G}\left(\frac{1}{\theta}\right)^{\otimes n}.$$

1. (5 points) Show that the posterior distribution is given by $\Pi[\cdot \mid \mathbf{X}] = IB(a+n, n\bar{X}_n - n + 2)$.
   Correction: We have

$$\begin{aligned}\pi(\theta|\mathbf{X}) \propto\ & \theta^{-a-2}(\theta-1)\prod_{i=1}^{n}\left(\left(1-\frac{1}{\theta}\right)^{X_i-1}\frac{1}{\theta}\right)\mathbb{1}_{(1,+\infty)}(\theta)\\ \propto\ & \theta^{-a-2}(\theta-1)\left(\frac{\theta-1}{\theta}\right)^{\sum X_i-n}\theta^{-n}\mathbb{1}_{(1,+\infty)}(\theta)\\ \propto\ & \theta^{-(a+n)-(n\bar{X}_n-n+2)}(\theta-1)^{(n\bar{X}_n-n+2)-1}\mathbb{1}_{(1,+\infty)}(\theta).\end{aligned}$$

   We recognize the general term of the distribution $IB(a+n, n\bar{X}_n - n + 2)$.

2. (4 points) Give the posterior expectation $m_{\mathbf{X}}$ and the posterior variance $v_{\mathbf{X}}$.
   Correction: We have from the previous question and the formula provided above

$$m_{\mathbf{X}} = \frac{a+n\bar{X}_n+1}{a+n-1} \quad \text{and} \quad v_{\mathbf{X}} = \frac{(a+n\bar{X}_n+1)(n\bar{X}_n-n+2)}{(a+n-1)^2(a+n-2)}$$

3. (7 points) Give a Bayes estimator, denoted T^*, for the prior Π and the loss $\ell(\theta, t) = \frac{(t-\theta)^2}{\theta(\theta-1)}$.
Correction: We proceed by minimizing the posterior risk, using the posterior from the first question. For an estimator T, we have

$$\begin{aligned}\rho(\Pi, T|\mathbf{X}) &= \int_1^{+\infty} \frac{(T(\mathbf{X})-\theta)^2}{\theta(\theta-1)}\pi(\theta|\mathbf{X})d\theta \\ &= \frac{1}{B(a+n, n\bar{X}_n - n + 2)}\int_1^{+\infty}(T(\mathbf{X})-\theta)^2\theta^{-a-n-1}(\theta-1)^{n\bar{X}_n - n}d\theta \\ &= \frac{B(a+n+2, n\bar{X}_n - n + 1)}{B(a+n, n\bar{X}_n - n + 2)}\int_1^{+\infty}(T(\mathbf{X})-\theta)^2 d\tilde{\Pi}_{\mathbf{X}}(\theta)\end{aligned}$$

where $\tilde{\Pi}_{\mathbf{X}} = IB(a+n+2, n\bar{X}_n - n + 1)$. This quantity is minimal for $T(\mathbf{X})$ given by the expectation of the distribution $\tilde{\Pi}_{\mathbf{X}}$, which is

$$T^*(\mathbf{X}) = \frac{a + n\bar{X}_n + 2}{a+n+1}.$$

4. (8 points) Calculate the posterior risk of T^*, and deduce the Bayes risk $\mathbf{R}_B(\Pi)$.
Correction: By noting that $\int_1^{+\infty}(T^*(\mathbf{X})-\theta)^2 d\tilde{\Pi}_{\mathbf{X}}(\theta)$ corresponds to the variance of the distribution $\tilde{\Pi}_{\mathbf{X}}$, and by using the relation $\Gamma(x+1) = x\Gamma(x)$, we have

$$\begin{aligned}\rho(\Pi, T^*|\mathbf{X}) &= \frac{B(a+n+2, n\bar{X}_n - n + 1)}{B(a+n, n\bar{X}_n - n + 2)}\frac{(a+n\bar{X}_n+2)(n\bar{X}_n - n + 1)}{(a+n+1)^2(a+n)} \\ &= \frac{\Gamma(a+n+2)\Gamma(n\bar{X}_n - n + 1)}{\Gamma(a+n\bar{X}_n+3)}\frac{\Gamma(a+n\bar{X}_n+2)}{\Gamma(a+n)\Gamma(n\bar{X}_n - n + 2)}\frac{(a+n\bar{X}_n+2)(n\bar{X}_n - n + 1)}{(a+n+1)^2(a+n)} \\ &= \frac{\Gamma(a+n\bar{X}_n+2)}{\Gamma(a+n\bar{X}_n+3)}\frac{\Gamma(n\bar{X}_n - n + 1)}{\Gamma(n\bar{X}_n - n + 2)}\frac{\Gamma(a+n+2)}{\Gamma(a+n)}\frac{(a+n\bar{X}_n+2)(n\bar{X}_n - n + 1)}{(a+n+1)^2(a+n)} \\ &= \frac{1}{a+n\bar{X}_n+2}\frac{1}{n\bar{X}_n - n + 1}(a+n+1)(a+n)\frac{(a+n\bar{X}_n+2)(n\bar{X}_n - n + 1)}{(a+n+1)^2(a+n)} \\ &= \frac{1}{n+a+1}\end{aligned}$$

We thus have: $\mathbf{R}_B(\Pi) = \mathbb{E}[\rho(\Pi, T^*|\mathbf{X})] = \frac{1}{n+a+1}$.

5. We set $T'(\mathbf{X}) = \frac{n\bar{X}_n+1}{n+1}$.

(a) (5 points) For $\theta > 1$, calculate the pointwise risk $\mathbf{R}(\theta, T')$.

Correction: We have

$$\begin{aligned}
\mathbf{R}(\theta, T') &= \frac{1}{\theta(\theta-1)}\mathbb{E}_\theta\left[\left(\frac{n\bar{X}_n+1}{n+1}-\theta\right)^2\right] \\
&= \frac{1}{\theta(\theta-1)}\mathbb{E}_\theta\left[\left(\frac{n}{n+1}(\bar{X}_n-\theta)+\frac{1-\theta}{n+1}\right)^2\right] \\
&= \frac{1}{\theta(\theta-1)}\left(\left(\frac{n}{n+1}\right)^2 \mathrm{Var}_\theta(\bar{X}_n)+\frac{(\theta-1)^2}{(n+1)^2}\right) \\
&= \frac{1}{\theta(\theta-1)}\left(\left(\frac{n}{n+1}\right)^2 \frac{\theta(\theta-1)}{n}+\frac{(\theta-1)^2}{(n+1)^2}\right) \\
&= \frac{n}{(n+1)^2}+\frac{\theta-1}{\theta(n+1)^2} \\
&= \frac{1}{n+1}-\frac{1}{\theta(n+1)^2}
\end{aligned}$$

where we used that $\mathrm{Var}_\theta(\bar{X}_n) = \frac{1-\frac{1}{\theta}}{n(1/\theta)^2} = \frac{\theta(\theta-1)}{n}$.

(b) (3 points) Deduce the maximal risk of T', denoted $\mathbf{R}_{\max}(T')$.

Correction: We have $\mathbf{R}_{\max}(T') = \sup_{\theta>1}\mathbf{R}(\theta, T') = \frac{1}{n+1}$.

(c) (6 points) Show that T' is minimax.

Correction: We notice that $\lim_{a\to 0}\mathbf{R}_B(\Pi)$ corresponds to $\mathbf{R}_{\max}(T')$, when a tends to zero, and to the limit of the Bayes risk for the prior $IB(a,2)$ an By a result from the course, T' is minimax.

$$\lim_{a\to 0}\frac{1}{n+a+1} = \frac{1}{n+1} = \mathbf{R}_{\max}(T')$$

6. (7 points) We wish to test:

$$H_0 : \theta \leqslant 2 \quad \text{against} \quad H_1 : \theta > 2$$

Determine the Bayes test for the balanced loss function takes the form

$$\varphi^*(\mathbf{X}) = \mathbb{1}_{\{\overline{X}_n > c_{a,n}\}},$$

where $c_{a,n}$ is a constant depending on a and n to be determined.

Hint: *You may use the following property: The median of a Beta distribution Beta(α, β) is less than $\frac{1}{2}$ if and only if $\alpha < \beta$.*

Correction: The Bayes test for the 0-1 loss is given by:

$$\varphi^*(\mathbf{X}) = \mathbb{1}_{\{\pi(H_1|\mathbf{X}) > \pi(H_0|\mathbf{X})\}} = \mathbb{1}_{\{\pi(\theta > 2|\mathbf{X}) > 1/2\}}$$

Recall that the posterior distribution is $\theta \mid \mathbf{X} \sim IB(\alpha_n, \beta_n)$ with:

$$\alpha_n = a + n \quad \text{and} \quad \beta_n = n\overline{X}_n - n + 2$$

By definition of the Inverse-Beta distribution, the variable $Z = 1/\theta$ follows a Beta distribution:

$$Z \mid \mathbf{X} \sim \text{Beta}(\alpha_n, \beta_n)$$

The condition $\theta > 2$ is equivalent to $1/\theta < 1/2$, i.e., $Z < 1/2$. The test becomes:

$$\varphi^*(\mathbf{X}) = 1 \iff \pi(Z < 1/2 \mid \mathbf{X}) > 1/2$$

This inequality holds if and only if the median of Z is strictly less than $1/2$. Using the hint provided, this occurs if and only if $\alpha_n < \beta_n$:

$$\begin{aligned} \alpha_n < \beta_n &\iff a + n < n\overline{X}_n - n + 2 \\ &\iff n\overline{X}_n > a + 2n - 2 \\ &\iff \overline{X}_n > 2 + \frac{a-2}{n} \end{aligned}$$

Thus, the Bayes test is:

$$\varphi^*(\mathbf{X}) = \mathbb{1}_{\{\overline{X}_n > 2 + \frac{a-2}{n}\}}$$

3. Hierarchical Normal Models (39 points)

We consider the following model: $X_1, \dots, X_n$ are n independent variables with respective distributions $\mathcal{N}(\mu_i, 1)$. We suppose that the prior on the means takes the form

$$\pi(\mu_1, \dots, \mu_n \mid \xi, \tau) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\mu_i - \xi)^2}{2\tau^2}},$$

where ξ and τ are two hyperparameters.

1. (8 points) Determine the marginal likelihood of the data given the hyperparameters, denoted $p(\mathbf{X} \mid \xi, \tau)$.

Correction: To find the marginal likelihood given ξ and τ, we integrate out the parameters $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$ in the joint distribution

$$p_{\boldsymbol{\mu}}(X_1, \dots, X_n)\pi(\mu_1, \dots, \mu_n \mid \xi, \tau) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(X_i - \mu_i)^2}{2}} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\mu_i - \xi)^2}{2\tau^2}}$$

Since the pairs (X_i, μ_i) are independent given (ξ, τ), we can handle each observation separately.

First approach: The marginal distribution of a single X_i is the convolution of the Gaussian noise $\mathcal{N}(0, 1)$ and the Gaussian prior $\mathcal{N}(\xi, \tau^2)$. Thus, $X_i \mid \xi, \tau \sim \mathcal{N}(\xi, 1 + \tau^2)$. The joint marginal likelihood is:

$$p(\mathbf{X} \mid \xi, \tau) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(1 + \tau^2)}} \exp\left(-\frac{(X_i - \xi)^2}{2(1 + \tau^2)}\right)$$

Second approach:

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(X_i - \mu_i)^2}{2}} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\mu_i - \xi)^2}{2\tau^2}} d\mu_i &= \frac{1}{2\pi\tau} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[(X_i - \mu_i)^2 + \frac{(\mu_i - \xi)^2}{\tau^2}\right]\right) d\mu_i \\ &= \frac{1}{2\pi\tau} \int_{-\infty}^{\infty} \exp\left(-\frac{\tau^2 + 1}{2\tau^2}\left[\mu_i^2 - 2\mu_i \frac{X_i\tau^2 + \xi}{\tau^2 + 1} + \frac{X_i^2\tau^2 + \xi^2}{\tau^2 + 1}\right]\right) d\mu_i\end{aligned}$$

Let $m = \frac{X_i\tau^2 + \xi}{\tau^2 + 1}$ and complete the square for μ_i. The integral becomes:

$$\frac{1}{2\pi\tau} \exp\left(-\frac{1}{2}\left[X_i^2 + \frac{\xi^2}{\tau^2} - \frac{\tau^2 + 1}{\tau^2} m^2\right]\right) \int_{-\infty}^{\infty} \exp\left(-\frac{\tau^2 + 1}{2\tau^2}(\mu_i - m)^2\right) d\mu_i.$$

The exponent in the first exponential is

$$\begin{aligned}-\frac{1}{2}\left[\frac{(X_i^2\tau^2 + \xi^2)(\tau^2 + 1) - (X_i\tau^2 + \xi)^2}{\tau^2(\tau^2 + 1)}\right] &= -\frac{(X_i^2\tau^4 + X_i^2\tau^2 + \xi^2\tau^2 + \xi^2) - (X_i^2\tau^4 + 2X_i\xi\tau^2 + \xi^2)}{2\tau^2(\tau^2 + 1)} \\ &= -\frac{X_i^2\tau^2 - 2X_i\xi\tau^2 + \xi^2\tau^2}{2\tau^2(\tau^2 + 1)} \\ &= -\frac{\tau^2(X_i^2 - 2X_i\xi + \xi^2)}{2\tau^2(\tau^2 + 1)} = -\frac{(X_i - \xi)^2}{2(1 + \tau^2)}\end{aligned}$$

Then, we finally have

$$\begin{aligned}\frac{1}{2\pi\tau} \exp\left(-\frac{(X_i - \xi)^2}{2(1 + \tau^2)}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{\tau^2 + 1}{2\tau^2}(\mu_i - m)^2\right) d\mu_i &= \frac{1}{2\pi\tau} \exp\left(-\frac{(X_i - \xi)^2}{2(1 + \tau^2)}\right) \sqrt{\frac{2\pi\tau^2}{\tau^2 + 1}} \\ &= \frac{1}{\sqrt{2\pi(1 + \tau^2)}} \exp\left(-\frac{(X_i - \xi)^2}{2(1 + \tau^2)}\right)\end{aligned}$$

2. (8 points) We adopt a Hierarchical Bayes approach and put an improper prior $\pi(\xi, \tau) \propto 1$ on the hyperparameters. Write down the posterior density $\pi(\xi, \tau \mid \mathbf{X})$ (up to a normalizing constant). Under what condition on n is this posterior well-defined (proper)?

Correction: The posterior is proportional to likelihood $\times$ prior:

$$\pi(\xi, \tau \mid \mathbf{X}) \propto p(\mathbf{X} \mid \xi, \tau) \cdot 1 \propto (1+\tau^2)^{-n/2} \exp\left(-\frac{\sum(X_i - \xi)^2}{2(1+\tau^2)}\right)$$

Using the decomposition $\sum(X_i - \xi)^2 = S^2 + n(\overline{X}_n - \xi)^2$, where $S^2 = \sum(X_i - \overline{X}_n)^2$:

$$\pi(\xi, \tau \mid \mathbf{X}) \propto (1+\tau^2)^{-n/2} \exp\left(-\frac{S^2}{2(1+\tau^2)}\right) \exp\left(-\frac{n(\overline{X}_n - \xi)^2}{2(1+\tau^2)}\right)$$

To check propriety, we integrate over the hyperparameters.

(a) Integrating over ξ:

$$\int_{-\infty}^{\infty} \exp\left(-\frac{n(\xi - \overline{X}_n)^2}{2(1+\tau^2)}\right) d\xi = \sqrt{\frac{2\pi(1+\tau^2)}{n}} \propto (1+\tau^2)^{1/2}$$

(b) Integrating the result over τ (on $(0, \infty)$): The marginal for τ is proportional to

$$(1+\tau^2)^{-n/2}(1+\tau^2)^{1/2} \exp\left(-\frac{S^2}{2(1+\tau^2)}\right) = (1+\tau^2)^{-(n-1)/2} \exp\left(-\frac{S^2}{2(1+\tau^2)}\right)$$

For large τ, the integrand behaves like $\tau^{-(n-1)}$. The integral converges if $n - 1 > 1$, i.e., $n > 2$. Thus, the posterior is well-defined for $n \geqslant 3$.

We now wish to test the hypothesis $H_0 : \mu_1 = \cdots = \mu_n$ against its complementary H_1. We fix the hyperparameters ξ and τ.

3. (5 points) Give the definition of the Bayes Factor $B^{\pi}_{0/1}$ in terms of marginal likelihoods. **Correction:** The Bayes Factor $B^{\pi}_{0/1}$ is defined as the ratio of the marginal likelihood of the data under model H_0 to the marginal likelihood of the data under model H_1:

$$B^{\pi}_{0/1} = \frac{\int_{H_0} p_\theta(\mathbf{X}) d\pi_0(\theta)}{\int_{H_1} p_\theta(\mathbf{X}) d\pi_1(\theta)}$$

where π_0 and π_1 be the restrictions of the prior π to H_0 and H_1.

4. (10 points) We admit that under the constrained model H_0, the common mean μ follows the prior $\mu \sim \mathcal{N}(\xi, \tau^2)$, and the prior is unchanged under H_1. Prove that the Bayes Factor is

$$B^{\pi}_{0/1} = \sqrt{\frac{(1+\tau^2)^n}{1+n\tau^2}} \exp\left(-\frac{\tau^2}{2}\left[\frac{S^2}{1+\tau^2} - \frac{n(n-1)(\overline{X}_n-\xi)^2}{(1+\tau^2)(1+n\tau^2)}\right]\right)$$

Correction: **Denominator (H_1):** The marginal likelihood under H_1 corresponds to the unconstrained case calculated in Question 1:

$$(2\pi(1+\tau^2))^{-n/2} \exp\left(-\frac{\sum(X_i-\xi)^2}{2(1+\tau^2)}\right)$$

Numerator (H_0): Under H_0, we have $X_i = \mu + \epsilon_i$ with $\mu \sim \mathcal{N}(\xi, \tau^2)$. This implies that the vector $\mathbf{X}$ follows a multivariate normal distribution with mean vector $\xi\mathbf{1}$ and covariance matrix $\Sigma = I_n + \tau^2\mathbf{1}\mathbf{1}^T$. Alternatively, we can integrate out μ as before:

$$\begin{aligned}
&\int \left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{(X_i-\mu)^2}{2}}\right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\mu-\xi)^2}{2\tau^2}} d\mu \\
&= (2\pi)^{-n/2}(2\pi\tau^2)^{-1/2} \int \exp\left(-\frac{1}{2}\sum_{i=1}^n (X_i-\mu)^2 - \frac{(\mu-\xi)^2}{2\tau^2}\right) d\mu \\
&= (2\pi)^{-n/2}(2\pi\tau^2)^{-1/2} e^{-S^2/2} \int \exp\left(-\frac{1}{2}\left[n(\mu-\bar{X}_n)^2 + \frac{(\mu-\xi)^2}{\tau^2}\right]\right) d\mu
\end{aligned}$$

Let $Q(\mu) = n(\mu-\bar{X}_n)^2 + \frac{1}{\tau^2}(\mu-\xi)^2$. Expanding and grouping by μ

$$Q(\mu) = \mu^2\left(n + \frac{1}{\tau^2}\right) - 2\mu\left(n\bar{X}_n + \frac{\xi}{\tau^2}\right) + \left(n\bar{X}_n^2 + \frac{\xi^2}{\tau^2}\right)$$

Let $A = n + \frac{1}{\tau^2} = \frac{n\tau^2+1}{\tau^2}$ and $m = \frac{n\bar{X}_n + \xi/\tau^2}{A}$. Completing the square: $Q(\mu) = A(\mu-m)^2 + \left(n\bar{X}_n^2 + \frac{\xi^2}{\tau^2}\right) - Am^2$. The integral over μ yields $\sqrt{2\pi/A}$. The remaining exponent is:

$$\begin{aligned}
&\left(n\bar{X}_n^2 + \frac{\xi^2}{\tau^2}\right) - \frac{(n\bar{X}_n + \xi/\tau^2)^2}{n+1/\tau^2} \\
&= \frac{(n\bar{X}_n^2 + \xi^2/\tau^2)(n+1/\tau^2) - (n^2\bar{X}_n^2 + 2n\bar{X}_n\xi/\tau^2 + \xi^2/\tau^4)}{n+1/\tau^2} \\
&= \frac{n\bar{X}_n^2/\tau^2 + n\xi^2/\tau^2 - 2n\bar{X}_n\xi/\tau^2}{n+1/\tau^2} = \frac{n(\bar{X}_n-\xi)^2}{n\tau^2+1}
\end{aligned}$$

Combining terms:

$$\begin{aligned}
&(2\pi)^{-n/2}\frac{1}{\sqrt{2\pi\tau^2}}\sqrt{\frac{2\pi\tau^2}{n\tau^2+1}} \exp\left(-\frac{1}{2}\left[S^2 + \frac{n(\bar{X}_n-\xi)^2}{1+n\tau^2}\right]\right) \\
&= (2\pi)^{-n/2}(1+n\tau^2)^{-1/2} \exp\left(-\frac{1}{2}\left[S^2 + \frac{n(\bar{X}_n-\xi)^2}{1+n\tau^2}\right]\right)
\end{aligned}$$

Bayes Factor:

$$B^{\pi}_{0/1} = \frac{(1+n\tau^2)^{-1/2}}{(1+\tau^2)^{-n/2}} \frac{\exp\left(-\frac{1}{2}\left[S^2 + \frac{n(\overline{X}_n-\xi)^2}{1+n\tau^2}\right]\right)}{\exp\left(-\frac{1}{2}\frac{S^2+n(\overline{X}_n-\xi)^2}{1+\tau^2}\right)}$$

This simplifies to:

$$B^{\pi}_{0/1} = \sqrt{\frac{(1+\tau^2)^n}{1+n\tau^2}} \exp\left(-\frac{\tau^2}{2}\left[\frac{S^2}{1+\tau^2} - \frac{n(n-1)(\overline{X}_n - \xi)^2}{(1+\tau^2)(1+n\tau^2)}\right]\right)$$

5. (8 points) What is your decision if we set the hyperparameter $\tau^2 \to 0$ and we consider an unbalanced loss function where the cost of a Type I error (a_1) is strictly greater than the cost of a Type II error (a_0), $a_1 > a_0$?

Correction: Introducing π_0 the prior probability of H_0 (and $1 - \pi_0$ the prior probability of H_1). The Bayes decision rule for a generic hypothesis test H_0 vs H_1 with costs a_1 (false positive) and a_0 (false negative) is to reject H_0 if:

$$\pi(H_1 \mid \mathbf{X}) > \frac{a_1}{a_0 + a_1} \iff \frac{\pi(H_1 \mid \mathbf{X})}{\pi(H_0 \mid \mathbf{X})} > \frac{a_1}{a_0}$$

This is equivalent to:

$$B^{\pi}_{0/1} < \frac{\pi_0 a_0}{(1 - \pi_0) a_1}$$

Now we analyze the behavior of the Bayes Factor $B^{\pi}_{0/1}$ as $\tau^2 \to 0$. As $\tau^2 \to 0$:

- The term inside the exponential approaches 0 (since τ^2 is in the numerator). Thus $\exp(\dots) \to 1$.
- The other factor: $\sqrt{\frac{(1+0)^n}{1+0}} = \sqrt{1} = 1$.

Therefore, as $\tau^2 \to 0$, $B^{\pi}_{0/1} \to 1$. The decision condition becomes:

$$1 < \frac{\pi_0 a_0}{(1 - \pi_0) a_1} \iff (1 - \pi_0) a_1 < \pi_0 a_0$$

SCRAP PAPER – WILL NOT BE GRADED.

SCRAP PAPER – WILL NOT BE GRADED.

Distribution	PDF / PMF	CDF	Mean	Variance
Poisson(λ)	$\frac{\lambda^k}{k!}e^{-\lambda}$, $k \geqslant 0$	$e^{-\lambda}\sum_{i=0}^{k}\frac{\lambda^i}{i!}$	λ	λ
Uniform $U[a,b]$	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$
Gamma(k,θ)	$\frac{\theta^k x^{k-1}e^{-\theta x}}{\Gamma(k)}\mathbb{1}_{x>0}$	$\int_0^x \frac{\theta^k t^{k-1}e^{-\theta t}}{\Gamma(k)}dt$	k/θ	k/θ^2
Geometric $\mathcal{G}(p)$	$(1-p)^{k-1}p$, $k \geqslant 1$	$1-(1-p)^k$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Normal $N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/(2\sigma^2)}$	$\int_{-\infty}^x f(t)dt$	μ	σ^2
Standard Normal $N(0,1)$	$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$	$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-t^2/2}dt$	0	1
Exponential $\exp(\lambda)$	$\lambda e^{-\lambda x}$, $x \geqslant 0$	$1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Binomial $Bin(n,p)$	$\binom{n}{k}p^k(1-p)^{n-k}$	$\sum_{i=0}^{k}\binom{n}{i}p^i(1-p)^{n-i}$	np	$np(1-p)$
Beta(a,b)	$\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$, $0 \leqslant x \leqslant 1$	$\int_0^x \frac{t^{a-1}(1-t)^{b-1}}{B(a,b)}dt$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$

Table 1: Summary of Common Probability Distributions