Chapter 11 Solutions

Andrew Wu
Wasserman: All of Statistics

June 30, 2025

Problem 11.1. Let $X_1, \ldots, X_n \sim N(\theta, \sigma^2)$. Let us assume that σ is known, for simplicity. Suppose we take as a prior $\theta \sim N(a, b^2)$. Show that the posterior for θ is

$$\theta | X^n \sim N(\overline{\theta}, \tau^2)$$

where

$$\overline{\theta} = w\overline{X} + (1-w)a, \tag{1}$$

$$w = \frac{1/\mathrm{se}^2}{1/\mathrm{se}^2 + 1/b^2}, \ \frac{1}{\tau^2} = \frac{1}{\mathrm{se}^2} + \frac{1}{b^2}. \tag{2}$$

and $\mathrm{se} = \sigma/\sqrt{n}$ is the standard error of the MLE $\overline{X}$.

Solution. By Bayes' theorem, we have

$$f(\theta|x^n) \propto f(x^n|\theta) f(\theta).$$

With the prior $\theta \sim N(a, b^2)$, we obtain

$$\begin{aligned}
f(\theta|x^n) &\propto \left[\prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x_i-\theta)^2\right)\right] \cdot \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{1}{2b^2}(\theta-a)^2\right) \\
&\propto \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot \frac{1}{b\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2b^2}(\theta-a)^2 + \sum_{i=1}^n -\frac{1}{2\sigma^2}(x_i-\theta)^2\right) \\
&\propto \exp\left(-\frac{1}{2b^2}(\theta-a)^2 - \sum_{i=1}^n \frac{1}{2\sigma^2}(x_i-\theta)^2\right) \\
&\propto \exp\left(-\frac{\theta^2}{2b^2} + \frac{a\theta}{b^2} - \frac{a^2}{2b^2} - \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2} - \frac{x_i\theta}{\sigma^2} + \frac{\theta^2}{2\sigma^2}\right)\right) \\
&\propto \exp\left(-\frac{\theta^2}{2b^2} + \frac{a\theta}{b^2} - \frac{a^2}{2b^2} - \frac{\sum_{i=1}^n x_i^2}{2\sigma^2} + \frac{n\theta\overline{X}}{\sigma^2} - \frac{n\theta^2}{2\sigma^2}\right) \\
&\propto \exp\left[-\left(\left(\frac{1}{2b^2} + \frac{n}{2\sigma^2}\right)\theta^2 - \left(\frac{a}{b^2} + \frac{n\overline{X}}{\sigma^2}\right)\theta + \frac{a^2}{2b^2} + \frac{\sum_{i=1}^n x_i^2}{2\sigma^2}\right)\right] \\
&\propto \exp\left[-\left(\theta\sqrt{\frac{1}{2b^2} + \frac{n}{2\sigma^2}} - \frac{\frac{a}{2b^2} + \frac{n\overline{X}}{2\sigma^2}}{\sqrt{\frac{1}{2b^2} + \frac{n}{2\sigma^2}}}\right)^2\right]
\end{aligned}$$

where we've been dropping terms without θ in them.

Observe that

$$\begin{aligned}\frac{1}{2\tau^2} &= \frac{1}{2\mathrm{se}^2} + \frac{1}{2b^2} \\ &= \frac{1}{2\tau^2/n} + \frac{1}{2b^2} \\ &= \frac{n}{2\sigma^2} + \frac{1}{2b^2}\end{aligned}$$

and

$$\begin{aligned}\overline{\theta} &= w\overline{X} + (1-w)a \\ &= \frac{\overline{X}/\mathrm{se}^2}{\frac{1}{\mathrm{se}^2} + \frac{1}{b^2}} + \frac{a/b^2}{\frac{1}{\mathrm{se}^2} + \frac{1}{b^2}} \\ &= \frac{n\overline{X}/\sigma^2 + a/b^2}{\frac{n}{\sigma^2} + \frac{1}{b^2}}.\end{aligned}$$

That means that

$$f(\theta|x^n) \propto \exp\left(-\frac{1}{2\tau^2}\left(\theta - \overline{\theta}\right)^2\right)$$

so $\theta|X^n \sim N(\overline{\theta}, \tau^2)$, as desired. □

Problem 11.3. Let $X_1, \ldots, X_n \sim \mathrm{Uniform}(0, \theta)$. Let $f(\theta) \propto 1/\theta$. Find the posterior density.

Solution. By Bayes' theorem, we have

$$f(\theta|x^n) \propto f(x^n|\theta) \cdot f(\theta).$$

Thus

$$f(\theta|x^n) \propto \left(\frac{1}{\theta}\right)^n \cdot \frac{1}{\theta} = \frac{1}{\theta^{n+1}}$$

when $\theta \geq \max(x_1, \ldots, x_n)$ and $f(\theta|x^n) = 0$ otherwise.

Let $x_{(n)} = \max(x_1, \ldots, x_n)$. Then, to find the posterior density, we compute

$$\begin{aligned}\int_{x_{(n)}}^{\infty} \frac{1}{\theta^{n+1}} d\theta &= \left[-\frac{1}{n\theta^n}\right]_{x_{(n)}}^{\infty} \\ &= \frac{1}{nx_{(n)}^n}\end{aligned}$$

and it follows that the posterior density is

$$f(\theta|x^n) = nx_{(n)}^n \cdot \frac{1}{\theta^{n+1}}$$

for $\theta \geq x_{(n)}$, and 0 otherwise. □

Problem 11.4. Suppose that 50 people are given a placebo and 50 are given a new treatment. 30 placebo patients show improvement while 40 treated patients show improvement. Let $\tau = p_2 - p_1$, where p_2 is the probability of improving under treatment and p_1 is the probability of improving under placebo.

Find the MLE of τ. Find the standard error and 90 percent confidence interval using the delta method.

Then let $\psi = \log\left(\left(\frac{p_1}{1-p_1}\right) \div \left(\frac{p_2}{1-p_2}\right)\right)$ be the log-odds ratio. Find the MLE of ψ. Use the delta method to find a 90 percent confidence interval for ψ.

Solution. Let $\widehat{p}_1, \widehat{p}_2$ be the MLEs of p_1, p_2 such that $\widehat{\tau} = \widehat{p}_2 - \widehat{p}_1$.

Let n be the total number of people, and let s_1 be the number of people who show improvement under the placebo. We have

$$\begin{aligned}\mathcal{L}(p_1) &= \prod_{i=1}^{n} f(X_i; p_1) \\ &= p_1^{s_1}(1-p_1)^{n-s_1}\end{aligned}$$

so

$$\begin{aligned}\ell(p_1) &= \log\left(p_1^{s_1}(1-p_1)^{n-s_1}\right) \\ &= s_1 \log p_1 + (n-s_1)\log(1-p_1)\end{aligned}$$

so

$$\ell'(p_1) = \frac{s_1}{p_1} - \frac{n-s_1}{1-p_1}.$$

Setting the log-likelihood equal to 0 yields $\widehat{p}_1 = \frac{s_1}{n}$. Using similar logic for $\widehat{p}_2$, we obtain

$$\begin{aligned}\widehat{\tau} &= \widehat{p}_2 - \widehat{p}_1 \\ &= 0.8 - 0.6 = 0.2.\end{aligned}$$

The standard error is the square root of the variance of the estimator; that is,

$$\begin{aligned}\mathrm{se}^2 &= \mathbb{V}\left(\frac{s_2}{n} - \frac{s_1}{n}\right) \\ &= \mathbb{V}\left(\frac{s_2}{n}\right) + \mathbb{V}\left(\frac{s_1}{n}\right) \\ &= \frac{1}{n^2}\left(\mathbb{V}(s_2) + \mathbb{V}(s_1)\right).\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{V}(s_1) &= \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) \\ &= n\mathbb{V}(X_i) \\ &= np_1(1-p_1)\end{aligned}$$

and it follows that

$$\begin{aligned}\mathrm{se} &= \sqrt{\frac{1}{n^2}(np_1(1-p_1) + np_2(1-p_2))} \\ &= \frac{1}{n}\sqrt{np_1(1-p_1) + np_2(1-p_2)}.\end{aligned}$$

As we don't have the true values of p_1 and p_2, we have

$$\begin{aligned}\widehat{\mathrm{se}} &= \frac{1}{50}\sqrt{50(0.6 \cdot 0.4) + 50(0.8 \cdot 0.2)} \\ &\approx 0.0894.\end{aligned}$$

Then, a 90 percent confidence interval is given by

$$(\widehat{\tau} - \widehat{\mathrm{se}} \cdot z_{\alpha/2}, \widehat{\tau} + \widehat{\mathrm{se}} \cdot z_{\alpha/2})$$

with $\alpha = 0.1$.

Regarding the log-odds ratio, we have

$$
\begin{aligned}
\widehat{\psi} &= \log\left(\left(\frac{\widehat{p}_1}{1-\widehat{p}_1}\right) \div \left(\frac{\widehat{p}_2}{1-\widehat{p}_2}\right)\right) \\
&= \log\left(\frac{s_1/n}{1-s_1/n} \cdot \frac{1-s_2/n}{s_2/n}\right) \\
&= \log\left(\frac{ns_1 - s_1 s_2}{ns_2 - s_1 s_2}\right) \\
&= \log\frac{300}{800} \approx -0.98.
\end{aligned}
$$

The multiparameter delta method gives us that

$$
\frac{\widehat{\psi}-\psi}{\widehat{\mathrm{se}}(\widehat{\psi})} \xrightarrow{\mathrm{d}} N(0,1)
$$

where

$$
\widehat{\mathrm{se}}(\widehat{\psi}) = \sqrt{(\widehat{\nabla} g)^T \widehat{J}_n (\widehat{\nabla} g)}.
$$

To compute ∇g, we note that with $g(p_1, p_2) = \log\left(\left(\frac{p_1}{1-p_1}\right) \div \left(\frac{p_2}{1-p_2}\right)\right) = \log\frac{p_1}{1-p_1} - \log\frac{p_2}{1-p_2}$, we have

$$
\begin{aligned}
\nabla g &= \begin{bmatrix} \frac{\partial g}{\partial p_1} \\ \frac{\partial g}{\partial p_2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{p_1(1-p_1)} \\ -\frac{1}{p_2(1-p_2)} \end{bmatrix}
\end{aligned}
$$

so

$$
\begin{aligned}
\widehat{\nabla} g &= \begin{bmatrix} \frac{1}{\widehat{p}_1(1-\widehat{p}_1)} \\ -\frac{1}{\widehat{p}_2(1-\widehat{p}_2)} \end{bmatrix} \\
&= \begin{bmatrix} \frac{25}{6} \\ -\frac{25}{4} \end{bmatrix}.
\end{aligned}
$$

Next, we need to compute the Fisher information. With $\ell_n(p_1, p_2) = \log[p_1^{s_1}(1-p_1)^{n-s_1} p_2^{s_2}(1-p_2)^{n-s_2}] = s_1 \log p_1 + (n-s_1)\log(1-p_1) + s_2 \log p_2 + (n-s_2)\log(1-p_2)$, we have

$$
\begin{aligned}
I_n(p_1, p_2) &= -\begin{bmatrix} \mathbb{E}_{p_1,p_2}\left(\frac{\partial^2 \ell_n}{\partial p_1^2}\right) & \mathbb{E}_{p_1,p_2}\left(\frac{\partial^2 \ell_n}{\partial p_1 \partial p_2}\right) \\ \mathbb{E}_{p_1,p_2}\left(\frac{\partial^2 \ell_n}{\partial p_2 \partial p_1}\right) & \mathbb{E}_{p_1,p_2}\left(\frac{\partial^2 \ell_n}{\partial p_2^2}\right) \end{bmatrix} \\
&= -\begin{bmatrix} \mathbb{E}\left(-\frac{s_1}{p_1^2} - \frac{n-s_1}{(1-p_1)^2}\right) & 0 \\ 0 & \mathbb{E}\left(-\frac{s_2}{p_2^2} - \frac{n-s_2}{(1-p_2)^2}\right) \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{p_1^2}\mathbb{E}(s_1) + \frac{1}{(1-p_1)^2}\mathbb{E}(n-s_1) & 0 \\ 0 & \frac{1}{p_2^2}\mathbb{E}(s_2) + \frac{1}{(1-p_2)^2}\mathbb{E}(n-s_2) \end{bmatrix} \\
&= \begin{bmatrix} \frac{n}{p_1(1-p_1)} & 0 \\ 0 & \frac{n}{p_2(1-p_2)} \end{bmatrix}.
\end{aligned}
$$

That means that

$$
\begin{aligned}
J_n(\widehat{p}_1, \widehat{p}_2) &= \begin{bmatrix} \frac{\widehat{p}_1(1-\widehat{p}_1)}{n} & 0 \\ 0 & \frac{\widehat{p}_2(1-\widehat{p}_2)}{n} \end{bmatrix} \\
&= \begin{bmatrix} 0.0048 & 0 \\ 0 & 0.0032 \end{bmatrix}.
\end{aligned}
$$

This yields $\widehat{\text{se}}(\widehat{\psi}) = \sqrt{5/24} \approx 0.4564$. Therefore a 90 percent confidence interval for ψ is

$$(\widehat{\psi} - z_{\alpha/2}\widehat{\text{se}}, \widehat{\psi} + z_{\alpha/2}\widehat{\text{se}})$$

with $\alpha = 0.1$, or equivalently,

$$(-0.98 - 0.4564 \cdot 1.64, -0.98 + 0.4564 \cdot 1.64).$$

$\square$

Problem 11.6. Let $X_1, \ldots, X_n \sim \text{Pois}(\lambda)$.

(a) Let $\lambda \sim \text{Gamma}(\alpha, \beta)$ be the prior. Show that the posterior is also a Gamma. Find the posterior mean.

(b) Find the Jeffreys' prior. Find the posterior.

Solution. Note that the prior for λ has PDF $f(\lambda) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} e^{-\lambda/\beta}$. Let $s = \sum_{i=1}^n x_i$.

(a) By Bayes' theorem, we have

$$f(\lambda|x^n) \propto f(x^n|\lambda) f(\lambda).$$

So

$$\begin{aligned} f(\lambda|x^n) &\propto f(x^n|\lambda) f(\lambda) \\ &\propto \left(\prod_{i=1}^n e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!} \right) \cdot \frac{1}{\beta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} e^{-\lambda/\beta} \\ &\propto e^{-n\lambda} \cdot \lambda^s \cdot \lambda^{\alpha-1} \cdot e^{-\lambda/\beta} \\ &\propto \lambda^{s+\alpha-1} e^{-\lambda(n+\frac{1}{\beta})} \end{aligned}$$

and it follows that $\lambda|x^n \sim \text{Gamma}\left(s+\alpha, \frac{1}{n+\frac{1}{b}}\right)$. The posterior mean is $\frac{s+\alpha}{n+\frac{1}{\beta}}$.

(b) The Jeffreys' Prior is $f(\lambda) \propto \sqrt{I(\lambda)}$.

We have

$$\begin{aligned} I(\lambda) &= -\mathbb{E}_\lambda \left(\frac{\partial^2}{\partial \lambda^2} \log f(X; \lambda) \right) \\ &= -\mathbb{E}_\lambda \left(\frac{\partial^2}{\partial \lambda^2} (-\lambda + x \log \lambda - \log(x!)) \right) \\ &= -\mathbb{E}_\lambda \left(-\frac{x}{\lambda^2} \right) \\ &= \frac{1}{\lambda}. \end{aligned}$$

It follows that the Jeffreys' Prior is $f(\lambda) \propto \frac{1}{\sqrt{\lambda}}$. Again, we compute the posterior through Bayes' theorem. We have

$$\begin{aligned} f(\lambda|x^n) &\propto f(x^n|\lambda) \cdot f(\lambda) \\ &\propto \left(\prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \right) \cdot \frac{1}{\sqrt{\lambda}} \\ &\propto e^{-n\lambda} \lambda^{s-\frac{1}{2}} \end{aligned}$$

so therefore $\lambda|x^n \sim \Gamma(s+\frac{1}{2}, \frac{1}{n})$.

$\square$

Problem 11.7. The data consist of n IID triples

$$(X_1, R_1, Y_1), \ldots, (X_n, Y_n, R_n).$$

Let B be a finite but very large number, such as 100^{100}. Any realistic sample size n will be small compared to B. Let

$$\theta = (\theta_1, \ldots, \theta_B)$$

be a vector of unknown parameters such that $0 \le \theta_j \le 1$ for $1 \le j \le B$. Let

$$\xi = (\xi_1, \ldots, \xi_B)$$

be a vector of known numbers such that

$$0 < \delta \le \xi_j \le 1 - \delta < 1, \quad 1 \le j \le B$$

where δ is some small positive number. Each data point (X_i, R_i, Y_i) is drawn in the following way:

1. Draw X_i uniformly from $\{1, \ldots, B\}$.
2. Draw $R_i \sim \text{Bernoulli}(\xi_{X_i})$.
3. If $R_i = 1$, then draw $Y_i \sim \text{Bernoulli}(\theta_{X_i})$. Otherwise do not draw Y_i.

In this exercise, $R_i = 0$ can be thought of as meaning "missing." Our goal is to estimate $\psi = \mathbb{P}(Y_i = 1)$.

Define $\widehat{\psi} = \frac{1}{n}\sum_{i=1}^{n} \frac{R_i Y_i}{\xi_{X_i}}$. Show that

$$\mathbb{E}(\widehat{\psi}) = \psi, \quad \mathbb{V}(\widehat{\psi}) \le \frac{1}{n\delta^2}.$$

Solution. We have

$$\begin{aligned}
\mathbb{E}(\widehat{\psi}) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{R_i Y_i}{\xi_{X_i}}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{E}\left(\frac{R_i Y_i}{\xi_{X_i}}\right)\right].
\end{aligned}$$

Note that here ξ_{X_i} can't be taken out of the expectation because X_i itself is a random variable, and as such we treat $\frac{R_i Y_i}{\xi_{X_i}}$ in totality as a random variable.

To compute this expectation, we now use the Law of Total Expectation. Note that X_i is drawn uniformly from $\{1, \ldots, B\}$. We have

$$\begin{aligned}
\mathbb{E}\left(\frac{R_i Y_i}{\xi_{X_i}}\right) &= \sum_{j=1}^{B} \mathbb{E}\left[\left(\frac{R_i Y_i}{\xi_{X_i}}\middle| X_i = j\right) \cdot \mathbb{P}(X_i = j)\right] \\
&= \frac{1}{B}\sum_{j=1}^{B} \mathbb{E}\left(\frac{R_i Y_i}{\xi_{X_i}}\middle| X_i = j\right) \\
&= \frac{1}{B}\sum_{j=1}^{B} \frac{1}{\xi_j}\mathbb{E}(R_i Y_i | X_i = j)
\end{aligned}$$

where we note that as soon as we condition on $X_i = j$ we obtain $\frac{1}{\xi_{X_i}} = \frac{1}{\xi_j}$, a constant we can take out of the expectation.

Now, with $\mathbb{E}(R_iY_i|X_i = j)$ we know that if $X_i = j$ then $R_i \sim \text{Bernoulli}(\xi_j)$, and $Y_i \sim \text{Bernoulli}(\theta_j)$ if $R_i = 1$ and Y_i does not exist otherwise. Thus we get

$$\begin{aligned}\mathbb{E}\left(\frac{R_iY_i}{\xi_{X_i}}\right) &= \frac{1}{B}\sum_{j=1}^{B}\frac{1}{\xi_j}\cdot\xi_j\theta_j \\ &= \frac{1}{B}\sum_{j=1}^{B}\theta_j = \psi.\end{aligned}$$

Regarding the variance: we have

$$\begin{aligned}\mathbb{V}(\widehat{\psi}) &= \mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{R_iY_i}{\xi_{X_i}}\right) \\ &= \frac{1}{n^2}\mathbb{V}\left(\sum_{i=1}^{n}\frac{R_iY_i}{\xi_{X_i}}\right)\end{aligned}$$

So we want to show that $\mathbb{V}\left(\sum_{i=1}^{n}\frac{R_iY_i}{\xi_{X_i}}\right) \leq \frac{n}{\delta^2}$, or equivalently that $\mathbb{V}\left(\frac{R_iY_i}{\xi_{X_i}}\right) \leq \frac{1}{\delta^2}$. But note that $\frac{R_iY_i}{\xi_{X_i}} \in \{0, \frac{1}{\xi_{X_i}}\}$. Moreover, we know that $\frac{1}{1-\delta} \leq \frac{1}{\xi_{X_i}} \leq \frac{1}{\delta}$.

The variance is a measure of the expected squared differential from the mean. That means that if $\frac{R_iY_i}{\xi_{X_i}}$ has a range of at most $\frac{1}{\delta}$, between 0 and $\frac{1}{\delta}$, then the expected squared differential from the mean can be at maximum $(\frac{1}{\delta})^2$. The conclusion follows. □

Problem 11.8. Let $X \sim N(\mu, 1)$. Consider testing

$$H_0 : \mu = 0 \quad \text{versus} \quad H_1 : \mu \neq 0.$$

Take $\mathbb{P}(H_0) = \mathbb{P}(H_1) = \frac{1}{2}$. Let the prior for μ under H_1 be $\mu \sim N(0, b^2)$. Find an expression for $\mathbb{P}(H_0|X = x)$. Compare $\mathbb{P}(H_0|X = x)$ to the p-value of the Wald test. Do the comparison numerically for a variety of values of x and b.

Now repeat the problem using a sample of size n. You will see that the posterior probability of H_0 can be large even when the p-value is small, especially when n is large.

Solution. We have

$$\begin{aligned}\mathbb{P}(H_0|X = x) &= \frac{f(x|H_0)\cdot\mathbb{P}(H_0)}{f(x)} \\ &= \frac{f(x|H_0)\cdot\mathbb{P}(H_0)}{f(x|H_0)\cdot\mathbb{P}(H_0) + f(x|H_1)\cdot\mathbb{P}(H_1)} \\ &= \frac{\frac{1}{2}f(x|H_0)}{\frac{1}{2}f(x|H_0) + \frac{1}{2}f(x|H_1)} \\ &= \frac{f(x|H_0)}{f(x|H_0) + f(x|H_1)}.\end{aligned}$$

Moreover,

$$f(x|H_0) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$$

and

$$f(x|H_1) = \int_{-\infty}^{\infty} f(x|\mu)\cdot f(\mu)d\mu.$$

The second equation can be thought of as follows: we need to find $f(x|H_1)$, but H_1 is not a single $\mu = 0$-type hypothesis, so we must integrate over all possible values of μ; to do so, we must "sum" the products of the "probability" of getting μ from $N(0,b^2)$ (that is, $f(\mu)$) and the "probability" of getting x given the parameter μ (that is, $f(x|\mu)$.) Here we thus have

$$f(\mu) = \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{1}{2b^2} \cdot \mu^2\right)$$

and

$$f(x|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\mu)^2\right).$$

Thus

$$\begin{aligned}
f(x|H_1) &= \int_{-\infty}^{\infty} \frac{1}{2\pi b} \exp\left(-\frac{1}{2}\left((x-\mu)^2 + \frac{\mu^2}{b^2}\right)\right) d\mu \\
&= \frac{1}{2\pi b} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left((x-\mu)^2 + \frac{\mu^2}{b^2}\right)\right) d\mu \\
&= \frac{1}{2\pi b} \int_{-\infty}^{\infty} \exp\left(-\left(\left(\sqrt{\frac{1}{2}+\frac{1}{2b^2}} \cdot \mu - \frac{x}{2\sqrt{\frac{1}{2}+\frac{1}{2b^2}}}\right)^2 + x^2\left(\frac{1}{2} - \frac{1}{2+\frac{2}{b^2}}\right)\right)\right) d\mu \\
&= \frac{1}{2\pi b} \exp\left(-x^2\left(\frac{1}{2} - \frac{b^2}{2b^2+2}\right)\right) \int_{-\infty}^{\infty} \exp\left(-\left(\sqrt{\frac{1}{2}+\frac{1}{2b^2}} \cdot \mu - \frac{x}{2\sqrt{\frac{1}{2}+\frac{1}{2b^2}}}\right)^2\right) d\mu \\
&= \frac{1}{2\pi b} \exp\left(-x^2\left(\frac{1}{2} - \frac{b^2}{2b^2+2}\right)\right) \int_{-\infty}^{\infty} \exp(-u^2) \frac{du}{\sqrt{\frac{1}{2}+\frac{1}{2b^2}}} \\
&= \frac{\sqrt{\pi}}{2\pi b} \cdot \frac{1}{\sqrt{\frac{1}{2}+\frac{1}{2b^2}}} \cdot \exp\left(-x^2\left(\frac{1}{2} - \frac{b^2}{2b^2+2}\right)\right) \\
&= \frac{1}{\sqrt{2\pi(1+b^2)}} \exp\left(-\frac{x^2}{2(b^2+1)}\right).
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{P}(H_0|X=x) &= \frac{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)}{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) + \frac{1}{\sqrt{2\pi(1+b^2)}} \exp\left(-\frac{x^2}{2(b^2+1)}\right)} \\
&= \frac{1}{1 + \frac{1}{\sqrt{1+b^2}} \exp\left(\frac{x^2}{2}\left(1 - \frac{1}{b^2+1}\right)\right)}.
\end{aligned}$$

Meanwhile, the Wald test has us reject H_0 when $|W| > z_{\alpha/2}$, where $W = \frac{\widehat{\mu}}{\widehat{\text{se}}}$. Our estimator $\widehat{\mu}$ is just $\widehat{\mu} = x$. So we have

$$\begin{aligned}
w &= \frac{x}{\sqrt{\mathbb{V}(X)}} \\
&= x
\end{aligned}$$

and we reject H_0 when $|x| > z_{\alpha/2}$. That means the p-value for the Wald test is $\mathbb{P}(|Z| > |x|)$.

We will do the numerical comparison now.

Take $x = 1, b = 1$. Then $\mathbb{P}(H_0|X=x) \approx 0.524$ and the Wald p-value is 0.3174.

Take $x = 3, b = 1$. Then $\mathbb{P}(H_0|X=x) \approx 0.13$ and the Wald p-value is 0.0026.

Take $x = 1, b = 3$. Then $\mathbb{P}(H_0|X=x) \approx 0.668$ and the Wald p-value is 0.3174.

Note that the Wald p-value doesn't actually depend on b.

As for the problem with a sample of size n: we have

$$\begin{aligned}\mathbb{P}(H_0|x^n) &= \frac{f(x^n|H_0)\cdot\mathbb{P}(H_0)}{f(x^n)}\\ &= \frac{f(x^n|H_0)\cdot\mathbb{P}(H_0)}{f(x^n|H_0)\cdot\mathbb{P}(H_0)+f(x^n|H_1)\cdot\mathbb{P}(H_1)}\\ &= \frac{\frac{1}{2}f(x^n|H_0)}{\frac{1}{2}f(x^n|H_0)+\frac{1}{2}f(x^n|H_1)}\\ &= \frac{f(x^n|H_0)}{f(x^n|H_0)+f(x^n|H_1)}.\end{aligned}$$

We also have

$$f(x^n|H_0) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n x_i^2\right)$$

and

$$f(x|H_1) = \int_{-\infty}^{\infty} f(x^n|\mu)\cdot f(\mu)d\mu.$$

So using

$$f(\mu) = \frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{1}{2b^2}\cdot\mu^2\right)$$

and

$$f(x^n|\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2\right)$$

we get

$$f(x|H_1) = \frac{1}{b}\left(\frac{1}{\sqrt{2\pi}}\right)^{n+1}\int_{-\infty}^{\infty}\exp\left(-\frac{\mu^2}{2b^2}-\frac{1}{2}\sum_{i=1}^n(x_i-\mu)^2\right)d\mu$$

which comes out to be

$$\frac{1}{b}\left(\frac{1}{\sqrt{2\pi}}\right)^{n+1}\int_{-\infty}^{\infty}\exp\left(-\left(\left(\mu\sqrt{\frac{1}{2b^2}+\frac{n}{2}}-\frac{\sum_{i=1}^n x_i}{2\sqrt{\frac{1}{2b^2}+\frac{n}{2}}}\right)^2+\frac{\sum_{i=1}^n x_i^2}{2}-\left(\frac{\sum_{i=1}^n x_i}{2\sqrt{\frac{1}{2b^2}+\frac{n}{2}}}\right)^2\right)\right)d\mu$$

which then simplifies to

$$\frac{1}{b}\left(\frac{1}{\sqrt{2\pi}}\right)^{n+1}\frac{1}{\sqrt{\frac{1}{2b^2}+\frac{n}{2}}}\exp\left(\left(\frac{\sum_{i=1}^n x_i}{2\sqrt{\frac{1}{2b^2}+\frac{n}{2}}}\right)^2-\frac{\sum_{i=1}^n x_i^2}{2}\right)\sqrt{\pi}$$

after doing a u-substitution.

So thus

$$f(x|H_1) = \left(\frac{1}{\sqrt{2\pi}}\right)^n\cdot\frac{1}{\sqrt{1+nb^2}}\cdot\exp\left(-\frac{1}{2}\left[\left(\sum_{i=1}^n x_i^2\right)-\frac{b^2}{1+nb^2}\left(\sum_{i=1}^n x_i\right)^2\right]\right).$$

It follows that

$$\begin{aligned}\mathbb{P}(H_0|x^n) &= \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n\exp\left(-\frac{1}{2}\sum_{i=1}^n x_i^2\right)}{\left(\frac{1}{\sqrt{2\pi}}\right)^n\exp\left(-\frac{1}{2}\sum_{i=1}^n x_i^2\right)+\left(\frac{1}{\sqrt{2\pi}}\right)^n\cdot\frac{1}{\sqrt{1+nb^2}}\cdot\exp\left(-\frac{1}{2}\left[\left(\sum_{i=1}^n x_i^2\right)-\frac{b^2}{1+nb^2}\left(\sum_{i=1}^n x_i\right)^2\right]\right)}\\ &= \frac{1}{1+\frac{1}{\sqrt{1+nb^2}}\exp\left(\frac{1}{2}\left[\frac{b^2}{1+nb^2}\left(\sum_{i=1}^n x_i\right)^2\right]\right)}.\end{aligned}$$

Note that the Wald test doesn't care about b but that the posterior probability can grow extremely large when b is large, as expected. □