

STAT 545 – Final Exam Practice

Expanded, Beginner-Friendly Solutions

How to read this packet. These solutions are written for someone who is still learning the ideas. So in many places I do three things:

- explain *what* theorem or idea is being used,
 - explain *why* it applies here,
 - then carry out the algebra carefully step by step.
-

Problem 1

Let (X_t) be a Markov chain on the finite state space

$$S = \{0, 1, \dots, n\}$$

with reflecting boundaries and interior transition probabilities

$$P(i, i+1) = p, \quad P(i, i-1) = 1-p, \quad 1 \leq i \leq n-1,$$

where $0 < p < 1$. At the boundaries we have

$$P(0, 1) = 1, \quad P(n, n-1) = 1.$$

This chain moves along the line of states

$$0, 1, 2, \dots, n.$$

If the chain is in the middle, it can move one step right or one step left. At the endpoints it is forced back inward.

(a) Communicating classes and period

Step 1: Find the communicating classes. A **communicating class** is a collection of states that can all reach one another.

Here the transition picture is

$$0 \leftrightarrow 1 \leftrightarrow 2 \leftrightarrow \dots \leftrightarrow n.$$

Why do we have arrows in both directions between neighboring states? Because for every interior state i ,

$$P(i, i+1) = p > 0 \quad \text{and} \quad P(i, i-1) = 1-p > 0.$$

So if two states are neighbors, each can move directly to the other with positive probability.

Now take any two states $a, b \in S$. To get from a to b , we can move one step at a time along the line. Since each needed step has positive probability, there is a positive-probability path from a to b . The same is true from b back to a .

Therefore every state communicates with every other state, so the chain is **irreducible**.

Hence there is exactly **one communicating class**, namely

$$S = \{0, 1, \dots, n\}.$$

Step 2: Find the period. The **period** of a state i is

$$d(i) = \gcd\{t \geq 1 : P^t(i, i) > 0\}.$$

This means: look at all possible return times to state i , then take their greatest common divisor.

Notice that every move changes the state by exactly ± 1 . So every step changes parity:

- even state $\rightarrow$ odd state,
- odd state $\rightarrow$ even state.

That means:

- after an odd number of steps, you must be at opposite parity from where you started,
- so you *cannot* return to the same state in an odd number of steps.

But you *can* return in two steps. For example:

$$i \rightarrow i + 1 \rightarrow i \quad \text{or} \quad i \rightarrow i - 1 \rightarrow i.$$

At the endpoints,

$$0 \rightarrow 1 \rightarrow 0 \quad \text{and} \quad n \rightarrow n - 1 \rightarrow n$$

also both happen with positive probability.

So all return times are even, and 2 is one of them. Therefore

$$d(i) = 2 \quad \text{for every } i \in S.$$

Conclusion:

one communicating class S , period 2.

(b) Stationary distribution

A **stationary distribution** is a probability distribution π satisfying

$$\pi P = \pi.$$

This means that if the chain starts in distribution π , then one step later it still has distribution π .

Because the chain is finite and irreducible, we know from standard Markov chain theory that a unique stationary distribution exists.

Step 1: Use detailed balance. For birth-death chains like this, the easiest method is often the **detailed balance equations**

$$\pi(i)P(i, j) = \pi(j)P(j, i).$$

These say that, in equilibrium, the probability flow from i to j matches the flow from j to i .

For neighboring interior states i and $i + 1$,

$$\pi(i)P(i, i + 1) = \pi(i + 1)P(i + 1, i).$$

Substitute the transition probabilities:

$$\pi(i)p = \pi(i + 1)(1 - p).$$

Solve for $\pi(i + 1)$:

$$\pi(i + 1) = \frac{p}{1 - p} \pi(i).$$

Define

$$r = \frac{p}{1 - p}.$$

Then

$$\pi(i + 1) = r \pi(i).$$

This means each time we move one state to the right, the stationary probability is multiplied by r .

Step 2: Express all states in terms of one unknown. From

$$\pi(i + 1) = r \pi(i),$$

we get

$$\pi(2) = r \pi(1), \quad \pi(3) = r \pi(2) = r^2 \pi(1),$$

and in general

$$\pi(i) = \pi(1)r^{i-1}, \quad 1 \leq i \leq n - 1.$$

Now relate $\pi(1)$ to $\pi(0)$ using the left boundary:

$$\pi(0)P(0, 1) = \pi(1)P(1, 0).$$

Since

$$P(0, 1) = 1, \quad P(1, 0) = 1 - p,$$

we get

$$\pi(0) = \pi(1)(1 - p),$$

so

$$\pi(1) = \frac{\pi(0)}{1 - p}.$$

Substitute this into the interior formula:

$$\pi(i) = \frac{\pi(0)}{1 - p} r^{i-1}, \quad 1 \leq i \leq n - 1.$$

For the right endpoint n , use the edge relation

$$\pi(n-1)P(n-1, n) = \pi(n)P(n, n-1).$$

That is,

$$\pi(n-1)p = \pi(n) \cdot 1,$$

so

$$\pi(n) = p\pi(n-1).$$

Since

$$\pi(n-1) = \frac{\pi(0)}{1-p}r^{n-2},$$

we get

$$\pi(n) = p \cdot \frac{\pi(0)}{1-p}r^{n-2} = \pi(0)r^{n-1}.$$

Step 3: Normalize. All probabilities must add to 1:

$$\sum_{i=0}^n \pi(i) = 1.$$

So

$$\pi(0) + \sum_{i=1}^{n-1} \frac{\pi(0)}{1-p}r^{i-1} + \pi(0)r^{n-1} = 1.$$

This determines $\pi(0)$, and then every other $\pi(i)$ follows.

Final form:

$$\begin{aligned} \pi(0) &= c, \\ \pi(i) &= \frac{c}{1-p} \left(\frac{p}{1-p} \right)^{i-1}, \quad 1 \leq i \leq n-1, \\ \pi(n) &= c \left(\frac{p}{1-p} \right)^{n-1}, \end{aligned}$$

where c is chosen so that the probabilities sum to 1.

(c) Mean return times when $p = 2/3$ and $n = 4$

We now plug in

$$p = \frac{2}{3}, \quad n = 4.$$

Step 1: Compute the ratio r .

$$r = \frac{p}{1-p} = \frac{2/3}{1/3} = 2.$$

Step 2: Write the stationary probabilities up to a constant. Let

$$\pi(0) = c.$$

Then

$$\pi(1) = \frac{c}{1-p} = \frac{c}{1/3} = 3c.$$

Now move to the right using multiplication by $r = 2$:

$$\pi(2) = 2\pi(1) = 6c,$$

$$\pi(3) = 2\pi(2) = 12c.$$

For the last state,

$$\pi(4) = c r^3 = 8c.$$

So

$$\pi = (c, 3c, 6c, 12c, 8c).$$

Step 3: Normalize. Since probabilities must sum to 1,

$$c + 3c + 6c + 12c + 8c = 1.$$

So

$$30c = 1 \quad \implies \quad c = \frac{1}{30}.$$

Thus

$$\pi(0) = \frac{1}{30}, \quad \pi(1) = \frac{3}{30}, \quad \pi(2) = \frac{6}{30}, \quad \pi(3) = \frac{12}{30}, \quad \pi(4) = \frac{8}{30}.$$

Step 4: Use the return-time formula. For an irreducible positive recurrent Markov chain,

$$E_i[T_i] = \frac{1}{\pi(i)}.$$

This means the expected return time to state i is the reciprocal of its long-run stationary probability.

Therefore,

$$E_1[T_1] = \frac{1}{\pi(1)} = \frac{1}{3/30} = 10,$$

$$E_2[T_2] = \frac{1}{\pi(2)} = \frac{1}{6/30} = 5,$$

$$E_3[T_3] = \frac{1}{\pi(3)} = \frac{1}{12/30} = \frac{30}{12} = \frac{5}{2}.$$

Answer:

$$E_1[T_1] = 10, \quad E_2[T_2] = 5, \quad E_3[T_3] = \frac{5}{2}.$$

(d) Fraction of time at the endpoints

A key interpretation of the stationary distribution is:

$$\pi(j) = \text{long-run fraction of time spent in state } j.$$

So the long-run fraction of time spent at state 0 is

$$\pi(0) = \frac{1}{30},$$

and at state 4 is

$$\pi(4) = \frac{8}{30} = \frac{4}{15}.$$

Interpretation. Because $p = 2/3$, the chain prefers moving right. So in the long run it spends much more time near the right side than the left side.

Problem 2

Consider a Markov chain on $\mathbb{N} = \{0, 1, 2, \dots\}$ with transitions

$$P(i, i+1) = p, \quad P(i, i-1) = 1-p, \quad i \geq 1,$$

and

$$P(0, 1) = 1.$$

This is similar to Problem 1, except now the state space is infinite:

$$0, 1, 2, 3, \dots$$

So we must be more careful. On an infinite state space, a stationary distribution might exist, or it might not.

(a) Stationary distribution for $p < 1/2$

Assume

$$p < \frac{1}{2}.$$

That means moving left is more likely than moving right, so the chain tends to drift back toward 0.

Step 1: Use detailed balance. For neighboring states i and $i+1$ with $i \geq 1$,

$$\pi(i)P(i, i+1) = \pi(i+1)P(i+1, i).$$

So

$$\pi(i)p = \pi(i+1)(1-p).$$

Hence

$$\pi(i+1) = \frac{p}{1-p}\pi(i).$$

Let

$$r = \frac{p}{1-p}.$$

Then

$$\pi(i+1) = r\pi(i).$$

Since $p < 1/2$, we have $r < 1$. This matters because if the probabilities shrink geometrically, the infinite sum may converge.

Step 2: Express everything in terms of $\pi(0)$. At the boundary,

$$\pi(0)P(0, 1) = \pi(1)P(1, 0).$$

So

$$\pi(0) \cdot 1 = \pi(1)(1-p),$$

hence

$$\pi(1) = \frac{\pi(0)}{1-p}.$$

Now for $i \geq 1$,

$$\pi(i) = \pi(1)r^{i-1} = \frac{\pi(0)}{1-p}r^{i-1}.$$

Step 3: Normalize. A probability distribution must sum to 1:

$$1 = \pi(0) + \sum_{i=1}^{\infty} \pi(i) = \pi(0) + \sum_{i=1}^{\infty} \frac{\pi(0)}{1-p} r^{i-1}.$$

Factor out the constant:

$$1 = \pi(0) + \frac{\pi(0)}{1-p} \sum_{i=1}^{\infty} r^{i-1}.$$

Since $|r| < 1$, we can use the geometric series formula:

$$\sum_{i=1}^{\infty} r^{i-1} = \sum_{k=0}^{\infty} r^k = \frac{1}{1-r}.$$

So

$$1 = \pi(0) + \frac{\pi(0)}{1-p} \cdot \frac{1}{1-r}.$$

Now simplify $1-r$:

$$1-r = 1 - \frac{p}{1-p} = \frac{1-p-p}{1-p} = \frac{1-2p}{1-p}.$$

Therefore

$$\frac{1}{1-r} = \frac{1-p}{1-2p}.$$

Substitute:

$$1 = \pi(0) + \frac{\pi(0)}{1-p} \cdot \frac{1-p}{1-2p} = \pi(0) + \frac{\pi(0)}{1-2p}.$$

Factor out $\pi(0)$:

$$1 = \pi(0) \left(1 + \frac{1}{1-2p} \right).$$

Combine the terms:

$$1 + \frac{1}{1-2p} = \frac{1-2p+1}{1-2p} = \frac{2(1-p)}{1-2p}.$$

So

$$1 = \pi(0) \frac{2(1-p)}{1-2p},$$

which gives

$$\pi(0) = \frac{1-2p}{2(1-p)}.$$

Step 4: Write the full stationary distribution. Since

$$\pi(i) = \frac{\pi(0)}{1-p} r^{i-1},$$

we get

$$\pi(i) = \frac{1-2p}{2(1-p)^2} \left(\frac{p}{1-p} \right)^{i-1}, \quad i \geq 1.$$

Thus

$$\boxed{\pi(0) = \frac{1-2p}{2(1-p)}, \quad \pi(i) = \frac{1-2p}{2(1-p)^2} \left(\frac{p}{1-p} \right)^{i-1}, \quad i \geq 1.}$$

Step 5: Classify the chain. Because the chain is irreducible and we found a valid stationary distribution, it is **positive recurrent**.

(b) Limiting fraction of time at 0

For an irreducible positive recurrent Markov chain, the long-run fraction of time spent in state j equals $\pi(j)$.

So the limiting fraction of time at 0 is

$$\pi(0) = \frac{1 - 2p}{2(1 - p)}.$$

In indicator notation,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{\{X_t=0\}} = \frac{1 - 2p}{2(1 - p)} \quad \text{almost surely.}$$

(c) Case $p = 1/2$

Now suppose

$$p = \frac{1}{2}.$$

This is the **fair** case.

We define

$$f(i) = P(\tau_0 < \infty \mid X_0 = i),$$

where τ_0 is the first time the chain hits 0.

So $f(i)$ means: starting from state i , what is the probability that the chain ever reaches 0?

Step 1: Write the recursion. If $i \geq 1$, then conditioning on the first step gives

$$f(i) = \frac{1}{2}f(i+1) + \frac{1}{2}f(i-1).$$

Why? Because from i the chain moves to $i+1$ or $i-1$ with equal probability, and then the future hitting probability is $f(i+1)$ or $f(i-1)$.

Multiply by 2:

$$2f(i) = f(i+1) + f(i-1).$$

Rearrange:

$$f(i+1) - f(i) = f(i) - f(i-1).$$

So the successive differences are constant, which means $f(i)$ must be linear:

$$f(i) = a + bi.$$

Step 2: Use the fact that $f(i)$ is a probability. Since $f(i)$ is a probability,

$$0 \leq f(i) \leq 1 \quad \text{for all } i.$$

If $b \neq 0$, then the linear function $a + bi$ will eventually leave the interval $[0, 1]$ as $i \rightarrow \infty$. So we must have

$$b = 0.$$

Thus $f(i)$ is constant.

Since

$$f(0) = 1,$$

that constant must be 1:

$$f(i) = 1 \quad \text{for all } i.$$

So the chain hits 0 with probability 1 from every starting state. Therefore the chain is **recurrent**.

Step 3: Check for a stationary distribution. Detailed balance gives

$$\pi(i+1) = \frac{p}{1-p} \pi(i).$$

If $p = 1/2$, then

$$\pi(i+1) = \pi(i), \quad i \geq 1.$$

So all states $1, 2, 3, \dots$ must have the same stationary probability.

At the boundary,

$$\pi(0) \cdot 1 = \pi(1) \cdot \frac{1}{2},$$

so

$$\pi(1) = 2\pi(0).$$

Hence

$$\pi(i) = 2\pi(0), \quad i \geq 1.$$

But then

$$\sum_{i=0}^{\infty} \pi(i) = \pi(0) + \sum_{i=1}^{\infty} 2\pi(0),$$

which diverges unless $\pi(0) = 0$, and that would force all probabilities to be zero.

So no stationary distribution exists.

Step 4: Final classification. The chain is recurrent, but it has no stationary distribution. Therefore it is **null recurrent**.

(d) Case $p > 1/2$

Now suppose

$$p > \frac{1}{2}.$$

This means the chain prefers moving to the right, so we expect it may escape to infinity and fail to return to 0.

Again define

$$f(i) = P(\tau_0 < \infty \mid X_0 = i).$$

Step 1: Write the recursion. For $i \geq 1$,

$$f(i) = pf(i+1) + (1-p)f(i-1),$$

and

$$f(0) = 1.$$

Step 2: Try an exponential solution. A common method for linear recursions is to guess

$$f(i) = \lambda^i.$$

Substitute into the recursion:

$$\lambda^i = p\lambda^{i+1} + (1-p)\lambda^{i-1}.$$

Divide by λ^{i-1} :

$$\lambda = p\lambda^2 + (1-p).$$

Rearrange:

$$p\lambda^2 - \lambda + (1-p) = 0.$$

The roots are

$$\lambda_1 = 1, \quad \lambda_2 = \frac{1-p}{p}.$$

Let

$$\rho = \frac{1-p}{p}.$$

Since $p > 1/2$, we have

$$0 < \rho < 1.$$

So the general solution is

$$f(i) = A + B\rho^i.$$

Step 3: Use boundary behavior. If the chain starts farther and farther to the right, the probability of ever hitting 0 should go to 0:

$$\lim_{i \rightarrow \infty} f(i) = 0.$$

In the formula

$$f(i) = A + B\rho^i,$$

the term $\rho^i \rightarrow 0$, so the limit is A . Hence

$$A = 0.$$

So

$$f(i) = B\rho^i.$$

Now use

$$f(0) = 1.$$

Then

$$1 = B\rho^0 = B,$$

so

$$B = 1.$$

Therefore

$$\boxed{f(i) = \left(\frac{1-p}{p}\right)^i}.$$

Step 4: Classify the chain. For $i \geq 1$,

$$f(i) = \left(\frac{1-p}{p} \right)^i < 1.$$

So there is a positive probability that the chain *never* reaches 0.

Therefore the chain is **transient**.

Problem 3

Suppose

$$Y_i \mid \beta \sim \text{Bernoulli}(p_i), \quad p_i = \frac{e^{\beta x_i}}{1 + e^{\beta x_i}}, \quad i = 1, \dots, n,$$

and the prior on β is

$$\beta \sim N(0, \tau^2).$$

We want to study the posterior distribution of β and write down a Metropolis-Hastings algorithm.

(a) Posterior distribution up to proportionality

Step 1: Write the likelihood. For Bernoulli data,

$$P(Y_i = y_i \mid \beta) = p_i^{y_i} (1 - p_i)^{1 - y_i}.$$

Because the observations are conditionally independent given β , the full likelihood is

$$L(\beta) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1 - y_i}.$$

Now plug in

$$p_i = \frac{e^{\beta x_i}}{1 + e^{\beta x_i}}, \quad 1 - p_i = \frac{1}{1 + e^{\beta x_i}}.$$

So

$$L(\beta) = \prod_{i=1}^n \left(\frac{e^{\beta x_i}}{1 + e^{\beta x_i}} \right)^{y_i} \left(\frac{1}{1 + e^{\beta x_i}} \right)^{1 - y_i}.$$

Step 2: Write the prior density. Since

$$\beta \sim N(0, \tau^2),$$

its density is proportional to

$$\exp\left(-\frac{\beta^2}{2\tau^2}\right).$$

Step 3: Multiply likelihood and prior. Bayes' rule tells us

$$p(\beta \mid y, x) \propto L(\beta) p(\beta).$$

Therefore

$$p(\beta \mid y, x) \propto \left[\prod_{i=1}^n \left(\frac{e^{\beta x_i}}{1 + e^{\beta x_i}} \right)^{y_i} \left(\frac{1}{1 + e^{\beta x_i}} \right)^{1 - y_i} \right] \exp\left(-\frac{\beta^2}{2\tau^2}\right).$$

Step 4: Write the log-posterior. Taking logs is useful because products become sums:

$$\log p(\beta \mid y, x) = \sum_{i=1}^n \left[y_i \beta x_i - \log(1 + e^{\beta x_i}) \right] - \frac{\beta^2}{2\tau^2} + C,$$

where C is a constant that does not depend on β .

So

$$\log p(\beta \mid y, x) = \sum_{i=1}^n \left[y_i \beta x_i - \log(1 + e^{\beta x_i}) \right] - \frac{\beta^2}{2\tau^2} + C.$$

(b) Random-walk proposal $\beta^* \sim N(\beta^{(t)}, s^2)$

Step 1: Recall the MH acceptance probability. In Metropolis-Hastings, if the current state is $\beta^{(t)}$ and we propose β^* , then the acceptance probability is

$$\alpha(\beta^{(t)}, \beta^*) = \min \left\{ 1, \frac{p(\beta^* \mid y, x) q(\beta^{(t)} \mid \beta^*)}{p(\beta^{(t)} \mid y, x) q(\beta^* \mid \beta^{(t)})} \right\}.$$

Step 2: Use symmetry of the proposal. Here

$$q(\beta^* \mid \beta^{(t)}) = \phi(\beta^*; \beta^{(t)}, s^2),$$

which is symmetric:

$$q(\beta^* \mid \beta^{(t)}) = q(\beta^{(t)} \mid \beta^*).$$

So those proposal densities cancel.

Hence

$$\alpha(\beta^{(t)}, \beta^*) = \min \left\{ 1, \frac{p(\beta^* \mid y, x)}{p(\beta^{(t)} \mid y, x)} \right\}.$$

Step 3: Write the ratio in exponential form. Using the log-posterior,

$$\log \frac{p(\beta^* \mid y, x)}{p(\beta^{(t)} \mid y, x)} = \sum_{i=1}^n y_i x_i (\beta^* - \beta^{(t)}) - \sum_{i=1}^n \log \frac{1 + e^{\beta^* x_i}}{1 + e^{\beta^{(t)} x_i}} - \frac{(\beta^*)^2 - (\beta^{(t)})^2}{2\tau^2}.$$

Exponentiating gives

$$\alpha = \min \left\{ 1, \exp \left[\sum_{i=1}^n y_i x_i (\beta^* - \beta^{(t)}) - \sum_{i=1}^n \log \frac{1 + e^{\beta^* x_i}}{1 + e^{\beta^{(t)} x_i}} - \frac{(\beta^*)^2 - (\beta^{(t)})^2}{2\tau^2} \right] \right\}.$$

Step 4: State the algorithm in words. One Metropolis-Hastings iteration is:

1. Start with the current value $\beta^{(t)}$.
2. Propose a new value

$$\beta^* \sim N(\beta^{(t)}, s^2).$$

3. Compute the acceptance probability α above.
4. With probability α , accept the proposal and set

$$\beta^{(t+1)} = \beta^*.$$

5. Otherwise reject it and keep

$$\beta^{(t+1)} = \beta^{(t)}.$$

(c) Non-symmetric proposal $\beta^* \sim N(\beta^{(t)} + c, s^2)$

Now the proposal is no longer centered at the current point, so the proposal is *not* symmetric.

Step 1: Write the general MH formula again.

$$\alpha = \min \left\{ 1, \frac{p(\beta^* | y, x) q(\beta^{(t)} | \beta^*)}{p(\beta^{(t)} | y, x) q(\beta^* | \beta^{(t)})} \right\}.$$

Step 2: Write the two Gaussian proposal densities. Forward proposal:

$$q(\beta^* | \beta^{(t)}) = \phi(\beta^*; \beta^{(t)} + c, s^2).$$

Reverse proposal:

$$q(\beta^{(t)} | \beta^*) = \phi(\beta^{(t)}; \beta^* + c, s^2).$$

For a normal density, terms not involving the exponent cancel in the ratio, so

$$\frac{q(\beta^{(t)} | \beta^*)}{q(\beta^* | \beta^{(t)})} = \exp \left(\frac{2c(\beta^{(t)} - \beta^*)}{s^2} \right).$$

Step 3: Final acceptance probability. Therefore

$$\alpha = \min \left\{ 1, \frac{p(\beta^* | y, x)}{p(\beta^{(t)} | y, x)} \exp \left(\frac{2c(\beta^{(t)} - \beta^*)}{s^2} \right) \right\}.$$

Why is the extra factor needed? Because the proposal is biased in one direction. The MH correction factor

$$\frac{q(\beta^{(t)} | \beta^*)}{q(\beta^* | \beta^{(t)})}$$

adjusts for that asymmetry so that the chain still has the correct posterior as its stationary distribution.

(d) Interpreting MCMC output

Suppose the posterior summary is:

- posterior mean = 2.80,
- posterior sd = 0.50,
- 95% credible interval (1.90, 3.80),
- $P(\beta > 0) = 0.999$,
- acceptance rate = 0.33.

(i) Is the proposal variance well tuned? An acceptance rate around 0.33 is usually considered reasonable for a one-dimensional random-walk Metropolis algorithm. It is not too small and not too large.

So the proposal scale appears to be **well tuned**.

(ii) **Is the prior influential?** The prior is $N(0, 25)$, so its standard deviation is

$$\sqrt{25} = 5.$$

The posterior standard deviation is only

$$0.50.$$

That is much smaller, which means the data have sharply concentrated the posterior.

So the data are **much more informative than the prior**, and the prior is not dominating the result.

(iii) **Is there evidence that $\beta > 0$?** Yes, very strong evidence.

Why?

- The 95% credible interval (1.90, 3.80) is entirely above 0.
- Also $P(\beta > 0) = 0.999$, meaning the posterior places 99.9% probability above zero.

So the posterior strongly supports

$$\beta > 0.$$

(iv) **Interpret e^β .** In logistic regression, e^β is an **odds ratio** associated with a one-unit increase in x .

If the posterior mode is near $\beta \approx 3$, then the MAP estimate of e^β is approximately

$$e^3 \approx 20.1.$$

A 95% credible interval for e^β is obtained by exponentiating the credible interval for β :

$$(e^{1.9}, e^{3.8}) \approx (6.69, 44.7).$$

So the estimated odds multiplier is about

$$20,$$

with a plausible range from about 6.7 to 44.7.

Interpretation in words. A one-unit increase in x multiplies the odds of success by roughly 20.

Problem 4

Consider the two-component Gaussian mixture model

$$y_i \mid z_i, \mu_0, \mu_1, \sigma^2 \sim \begin{cases} N(\mu_0, \sigma^2), & z_i = 0, \\ N(\mu_1, \sigma^2), & z_i = 1, \end{cases}$$

with latent allocations

$$z_i \mid \pi \sim \text{Bernoulli}(\pi),$$

prior

$$\pi \sim \text{Beta}(a, b),$$

priors

$$\mu_0 \sim N(m_0, s_0^2), \quad \mu_1 \sim N(m_1, s_1^2),$$

and

$$\sigma^2 \sim \text{Inv-Gamma}(\alpha, \beta).$$

We derive the full conditional distributions for a Gibbs sampler.

(a) Full conditional for z_i

We want

$$P(z_i = 1 \mid \pi, \mu_0, \mu_1, \sigma^2, y_i).$$

Step 1: Use Bayes' rule. Conditioning on everything except z_i , the two possibilities are $z_i = 1$ or $z_i = 0$.

Up to proportionality,

$$P(z_i = 1 \mid \cdots) \propto P(z_i = 1 \mid \pi) f(y_i \mid z_i = 1, \mu_1, \sigma^2),$$

$$P(z_i = 0 \mid \cdots) \propto P(z_i = 0 \mid \pi) f(y_i \mid z_i = 0, \mu_0, \sigma^2).$$

Now,

$$P(z_i = 1 \mid \pi) = \pi, \quad P(z_i = 0 \mid \pi) = 1 - \pi,$$

and

$$f(y_i \mid z_i = 1, \mu_1, \sigma^2) = \phi(y_i; \mu_1, \sigma^2),$$

$$f(y_i \mid z_i = 0, \mu_0, \sigma^2) = \phi(y_i; \mu_0, \sigma^2).$$

Step 2: Normalize. Therefore

$$P(z_i = 1 \mid \cdots) = \frac{\pi \phi(y_i; \mu_1, \sigma^2)}{\pi \phi(y_i; \mu_1, \sigma^2) + (1 - \pi) \phi(y_i; \mu_0, \sigma^2)}.$$

So

$$z_i \mid \cdots \sim \text{Bernoulli}(w_i), \quad w_i = \frac{\pi \phi(y_i; \mu_1, \sigma^2)}{\pi \phi(y_i; \mu_1, \sigma^2) + (1 - \pi) \phi(y_i; \mu_0, \sigma^2)}.$$

Interpretation. Each observation is probabilistically assigned to one of the two mixture components. If y_i is more likely under component 1 than component 0, then w_i will be closer to 1.

(b) Full conditional for π

Let

$$n_1 = \sum_{i=1}^n z_i, \quad n_0 = n - n_1.$$

Step 1: Write prior and likelihood in terms of π . The prior is

$$\pi \sim \text{Beta}(a, b),$$

so its density is proportional to

$$\pi^{a-1}(1-\pi)^{b-1}.$$

Given the z_i values, the likelihood for π is

$$\prod_{i=1}^n \pi^{z_i} (1-\pi)^{1-z_i} = \pi^{n_1} (1-\pi)^{n_0}.$$

Step 2: Multiply prior and likelihood. Thus the conditional density is proportional to

$$\pi^{a-1}(1-\pi)^{b-1} \pi^{n_1} (1-\pi)^{n_0} = \pi^{a+n_1-1} (1-\pi)^{b+n_0-1}.$$

That is the kernel of a Beta distribution.

So

$$\boxed{\pi \mid z \sim \text{Beta}(a + n_1, b + n_0)}.$$

(c) Full conditionals for μ_0 and μ_1

We derive μ_0 ; the derivation for μ_1 is identical.

Let

$$I_0 = \{i : z_i = 0\}, \quad n_0 = |I_0|.$$

Step 1: Collect all terms involving μ_0 . The prior is

$$\mu_0 \sim N(m_0, s_0^2).$$

For each i with $z_i = 0$, we have

$$y_i \mid \mu_0, \sigma^2 \sim N(\mu_0, \sigma^2).$$

Normal prior + normal likelihood gives a normal posterior. So we know the full conditional will be normal.

Step 2: State the posterior variance and mean. The posterior precision is the sum of prior precision and data precision:

$$\frac{1}{v_0^*} = \frac{1}{s_0^2} + \frac{n_0}{\sigma^2}.$$

So

$$v_0^* = \left(\frac{1}{s_0^2} + \frac{n_0}{\sigma^2} \right)^{-1}.$$

The posterior mean is a precision-weighted average:

$$m_0^* = v_0^* \left(\frac{m_0}{s_0^2} + \frac{\sum_{i:z_i=0} y_i}{\sigma^2} \right).$$

Hence

$$\mu_0 \mid z, \sigma^2, y \sim N(m_0^*, v_0^*),$$

where

$$v_0^* = \left(\frac{1}{s_0^2} + \frac{n_0}{\sigma^2} \right)^{-1}, \quad m_0^* = v_0^* \left(\frac{m_0}{s_0^2} + \frac{\sum_{i:z_i=0} y_i}{\sigma^2} \right).$$

Similarly,

$$\mu_1 \mid z, \sigma^2, y \sim N(m_1^*, v_1^*),$$

with

$$v_1^* = \left(\frac{1}{s_1^2} + \frac{n_1}{\sigma^2} \right)^{-1}, \quad m_1^* = v_1^* \left(\frac{m_1}{s_1^2} + \frac{\sum_{i:z_i=1} y_i}{\sigma^2} \right).$$

Interpretation. Each posterior mean is a compromise between:

- the prior mean, and
- the sample mean of the data assigned to that component.

(d) Full conditional for σ^2

Step 1: Write the prior. We have

$$\sigma^2 \sim \text{Inv-Gamma}(\alpha, \beta).$$

Its kernel is

$$(\sigma^2)^{-(\alpha+1)} \exp\left(-\frac{\beta}{\sigma^2}\right).$$

Step 2: Write the likelihood in terms of σ^2 . Conditional on allocations and means, the observations are normal with variance σ^2 . So the likelihood kernel is

$$(\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu_{z_i})^2\right).$$

Step 3: Multiply prior and likelihood. Combining exponents and powers of σ^2 gives

$$(\sigma^2)^{-(\alpha+1+n/2)} \exp\left(-\frac{1}{\sigma^2} \left[\beta + \frac{1}{2} \sum_{i=1}^n (y_i - \mu_{z_i})^2 \right]\right).$$

That is the kernel of an inverse-gamma distribution.

So

$$\sigma^2 \mid z, \mu_0, \mu_1, y \sim \text{Inv-Gamma}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2} \sum_{i=1}^n (y_i - \mu_{z_i})^2\right).$$

(e) One Gibbs sampling sweep

A Gibbs sampler cycles through the full conditionals one at a time.

One full sweep is:

1. For each $i = 1, \dots, n$, sample

$$z_i \mid \pi, \mu_0, \mu_1, \sigma^2, y_i.$$

2. Sample

$$\pi \mid z.$$

3. Sample

$$\mu_0 \mid z, \sigma^2, y.$$

4. Sample

$$\mu_1 \mid z, \sigma^2, y.$$

5. Sample

$$\sigma^2 \mid z, \mu_0, \mu_1, y.$$

Main idea. Each step samples from the exact conditional distribution of one block of unknowns, given the current values of everything else.

Problem 5

Suppose meteor arrivals are modeled by a Poisson process.

(a) Off-peak homogeneous Poisson process

Assume arrivals occur as a homogeneous Poisson process with constant rate λ meteors per minute. Suppose 28 meteors were observed in 120 minutes.

(i) Find the MLE of λ

For a homogeneous Poisson process observed for time T , the count

$$N(T) \sim \text{Poisson}(\lambda T).$$

The maximum likelihood estimator is

$$\hat{\lambda} = \frac{N(T)}{T}.$$

Here

$$N(T) = 28, \quad T = 120.$$

So

$$\hat{\lambda} = \frac{28}{120} = \frac{7}{30} \approx 0.2333$$

meteors per minute.

$$\hat{\lambda} = \frac{7}{30} \approx 0.2333$$

(ii) Correct for 80% detection

If only 80% of meteors are detected, then the observed rate is only

$$0.8 \times (\text{true rate}).$$

So

$$\hat{\lambda}_{\text{true}} = \frac{\hat{\lambda}_{\text{obs}}}{0.8} = \frac{0.2333}{0.8} \approx 0.2917.$$

$$\hat{\lambda}_{\text{true}} \approx 0.2917$$

(iii) Expected extra time to see 2 more meteors

In a Poisson process with rate λ , the waiting time to the next event is exponential with mean

$$\frac{1}{\lambda}.$$

More generally, the waiting time to k additional events has Gamma distribution with mean

$$\frac{k}{\lambda}.$$

So the expected additional time to see 2 more meteors is

$$\frac{2}{\hat{\lambda}} = \frac{2}{7/30} = \frac{60}{7} \approx 8.57 \text{ minutes.}$$

$$\frac{60}{7} \approx 8.57 \text{ minutes}$$

(iv) Is a shortest gap of 3 seconds surprising?

First convert 3 seconds to minutes:

$$3 \text{ sec} = \frac{3}{60} = 0.05 \text{ min.}$$

For a homogeneous Poisson process, one interarrival time

$$W \sim \text{Exponential}(\lambda),$$

so

$$P(W \leq 0.05) = 1 - e^{-\lambda(0.05)}.$$

Using $\hat{\lambda} = 7/30$,

$$P(W \leq 0.05) = 1 - e^{-(7/30)(0.05)} \approx 0.0116.$$

So for *one particular gap*, a 3-second wait is somewhat rare.

But if there were 28 meteors, then there are about 27 interarrival gaps to examine. The more gaps we look at, the less surprising it becomes that *at least one* of them is very short.

Approximate:

$$P(\min W_i \leq 0.05) = 1 - P(W_1 > 0.05, \dots, W_{27} > 0.05).$$

Assuming independence of the interarrival times,

$$P(\min W_i \leq 0.05) = 1 - \left(e^{-\lambda(0.05)}\right)^{27} = 1 - e^{-27\lambda(0.05)}.$$

Plug in $\lambda = 7/30$:

$$1 - e^{-27(7/30)(0.05)} \approx 0.27.$$

So the chance is about 27%. That is *not* especially surprising.

A 3-second shortest gap is consistent with the homogeneous model.

(b) Nonhomogeneous Poisson process

Now suppose the rate changes with time:

$$\lambda(t) = \begin{cases} 0.20, & 0 \leq t \leq 60, \\ 0.20 + \frac{t-60}{30}, & 60 < t \leq 90, \\ 0.20 + \frac{120-t}{30}, & 90 < t \leq 120, \\ 0.20, & 120 < t \leq 180. \end{cases}$$

(i) Describe/sketch the rate function

From $t = 0$ to 60, the rate is constant at 0.2.

From 60 to 90, it increases linearly.

At $t = 90$,

$$\lambda(90) = 0.20 + \frac{90 - 60}{30} = 0.20 + 1 = 1.20.$$

From 90 to 120, it decreases linearly back down.

From 120 to 180, it stays constant again at 0.2.

So the graph is:

- flat at 0.2,
- then rising linearly to 1.2,
- then falling linearly back to 0.2,
- then flat again.

(ii) Expected meteors in the first hour

For a nonhomogeneous Poisson process, the expected number of arrivals on $[a, b]$ is

$$E[N(b) - N(a)] = \int_a^b \lambda(t) dt.$$

On $[0, 60]$, the rate is constant 0.2, so

$$E[N(60)] = \int_0^{60} 0.2 dt = 0.2(60) = 12.$$

12

(iii) Probability the first wait is at least 15 minutes

For an NHPP, the probability of no arrivals by time t is

$$P(T_1 > t) = \exp\left(-\int_0^t \lambda(u) du\right).$$

For $t = 15$, the rate is still constant at 0.2, so

$$\int_0^{15} \lambda(u) du = \int_0^{15} 0.2 du = 3.$$

Therefore

$$P(T_1 \geq 15) = e^{-3} \approx 0.0498.$$

$e^{-3} \approx 0.0498$

(iv) **Compare** $[75, 105]$ **to** $[0, 30]$

First interval:

$$E[N(30) - N(0)] = \int_0^{30} 0.2 \, dt = 6.$$

Now compute the expected count on $[75, 105]$:

$$\int_{75}^{105} \lambda(t) \, dt = \int_{75}^{90} \left(0.2 + \frac{t-60}{30} \right) dt + \int_{90}^{105} \left(0.2 + \frac{120-t}{30} \right) dt.$$

For the first piece,

$$\int_{75}^{90} \left(0.2 + \frac{t-60}{30} \right) dt = 6.75.$$

For the second piece,

$$\int_{90}^{105} \left(0.2 + \frac{120-t}{30} \right) dt = 6.75.$$

So total expected count on $[75, 105]$ is

$$13.5.$$

Difference:

$$13.5 - 6 = 7.5.$$

$$\boxed{7.5}$$

Interpretation. The peak window $[75, 105]$ is expected to have 7.5 more meteors than the early baseline window $[0, 30]$.

Problem 6

Suppose a 2D lattice random walk takes 16 steps with counts

$$N_E = 6, \quad N_W = 2, \quad N_N = 5, \quad N_S = 3,$$

so the total number of steps is

$$n = 16.$$

We assume each step is one of East, West, North, South with probabilities

$$p_E, p_W, p_N, p_S, \quad p_E + p_W + p_N + p_S = 1.$$

(a) MLEs of the step probabilities

This is a multinomial model. The maximum likelihood estimator for each category probability is

$$\hat{p}_j = \frac{\text{count in category } j}{n}.$$

Therefore

$$\hat{p}_E = \frac{6}{16}, \quad \hat{p}_W = \frac{2}{16}, \quad \hat{p}_N = \frac{5}{16}, \quad \hat{p}_S = \frac{3}{16}.$$

So

$$\boxed{\hat{p}_E = \frac{6}{16}, \quad \hat{p}_W = \frac{2}{16}, \quad \hat{p}_N = \frac{5}{16}, \quad \hat{p}_S = \frac{3}{16}.$$

(b) One-step drift vector

The **drift** is the expected change in position in one step.

Let one step be written as the vector

$$\Delta Z_k = \begin{cases} (1, 0), & \text{East,} \\ (-1, 0), & \text{West,} \\ (0, 1), & \text{North,} \\ (0, -1), & \text{South.} \end{cases}$$

Then

$$E[\Delta Z_k] = (1, 0)p_E + (-1, 0)p_W + (0, 1)p_N + (0, -1)p_S.$$

Combine the x and y parts:

$$E[\Delta Z_k] = (p_E - p_W, p_N - p_S).$$

Using the MLEs,

$$\hat{\delta} = \left(\frac{6}{16} - \frac{2}{16}, \frac{5}{16} - \frac{3}{16} \right) = \left(\frac{1}{4}, \frac{1}{8} \right).$$

$$\boxed{\hat{\delta} = \left(\frac{1}{4}, \frac{1}{8} \right)}$$

(c) Sample mean increment vector

The average observed increment is

$$\bar{\Delta} = \frac{1}{n} \sum_{k=1}^n \Delta Z_k.$$

The total horizontal displacement is

$$N_E - N_W = 6 - 2 = 4,$$

and the total vertical displacement is

$$N_N - N_S = 5 - 3 = 2.$$

So

$$\bar{\Delta} = \left(\frac{4}{16}, \frac{2}{16} \right) = \left(\frac{1}{4}, \frac{1}{8} \right).$$

Thus

$$\boxed{\bar{\Delta} = \left(\frac{1}{4}, \frac{1}{8} \right)}.$$

Why does this match part (b)? Because the MLE for the drift is just the empirical average step.

(d) Brownian motion with drift, sampled every 2 seconds

Now interpret the same path as observations from a Brownian motion with drift:

$$X_t = \mu_x t + \text{noise}, \quad Y_t = \mu_y t + \text{noise},$$

sampled every 2 seconds.

With 16 steps, the total elapsed time is

$$16 \times 2 = 32 \text{ seconds}.$$

The net displacement is still

$$(4, 2).$$

For Brownian motion with drift, the MLE of the drift vector is net displacement divided by total time:

$$\hat{\mu} = \left(\frac{4}{32}, \frac{2}{32} \right) = \left(\frac{1}{8}, \frac{1}{16} \right).$$

So

$$\boxed{\hat{\mu}_x = \frac{1}{8}, \quad \hat{\mu}_y = \frac{1}{16}}.$$

(e) Compare the two drifts

From the random-walk model, the drift per step was

$$\left(\frac{1}{4}, \frac{1}{8}\right).$$

From the Brownian-time model, the drift per second is

$$\left(\frac{1}{8}, \frac{1}{16}\right).$$

These are not numerically identical because they use different units:

- part (b) gives expected displacement *per step*,
- part (d) gives expected displacement *per second*.

Since each step takes 2 seconds, dividing the per-step drift by 2 gives

$$\left(\frac{1}{4}, \frac{1}{8}\right) \div 2 = \left(\frac{1}{8}, \frac{1}{16}\right),$$

which matches the Brownian estimate.

Conclusion. The two answers are consistent once we account for the time scale.

Problem 7

This problem uses a plotted 2D trajectory and asks us to approximate a discrete random-walk model and a continuous-time Brownian-drift model from the figure.

Because the exact coordinates come from reading the plot, the answers are approximate.

Suppose the visible trajectory is approximately

$$(0, 0) \rightarrow (1.5, 1) \rightarrow (2.75, 2) \rightarrow (4, 2.75) \rightarrow (4.75, 4.25) \rightarrow (5.75, 3.5).$$

Then the approximate increments are

$$(1.5, 1), \quad (1.25, 1), \quad (1.25, 0.75), \quad (0.75, 1.5), \quad (1, -0.75).$$

(a) Classify each increment

We classify each increment by its dominant direction:

- $(1.5, 1)$ is mostly East,
- $(1.25, 1)$ is mostly East,
- $(1.25, 0.75)$ is mostly East,
- $(0.75, 1.5)$ is mostly North,
- $(1, -0.75)$ is mostly East.

So the approximate sequence is

$$E, \ E, \ E, \ N, \ E.$$

Hence the counts are approximately

$$N_E = 4, \quad N_W = 0, \quad N_N = 1, \quad N_S = 0.$$

$$(N_E, N_W, N_N, N_S) \approx (4, 0, 1, 0)$$

(b) MLEs of the directional probabilities

There are about 5 increments total, so the multinomial MLEs are

$$\hat{p}_E = \frac{4}{5}, \quad \hat{p}_W = 0, \quad \hat{p}_N = \frac{1}{5}, \quad \hat{p}_S = 0.$$

$$\hat{p}_E \approx \frac{4}{5}, \quad \hat{p}_W \approx 0, \quad \hat{p}_N \approx \frac{1}{5}, \quad \hat{p}_S \approx 0$$

(c) Drift under the random-walk model

For the discrete random-walk model, the drift vector is

$$\delta = (p_E - p_W, \ p_N - p_S).$$

Plugging in the estimates:

$$\hat{\delta} = \left(\frac{4}{5} - 0, \ \frac{1}{5} - 0 \right) = \left(\frac{4}{5}, \ \frac{1}{5} \right).$$

$$\hat{\delta} \approx \left(\frac{4}{5}, \ \frac{1}{5} \right)$$

(d) Drift under the Brownian-motion model

Now approximate the path as a Brownian motion with drift observed at uneven times.

Suppose the time increments are approximately

$$2, 2.25, 1.75, 2, 2.25 \text{ seconds,}$$

so the total time is

$$2 + 2.25 + 1.75 + 2 + 2.25 = 10.25 \text{ seconds.}$$

The net displacement is approximately

$$(5.75, 3.5).$$

For Brownian motion with drift, the drift estimate is

$$\hat{\mu} = \frac{\text{net displacement}}{\text{total time}} = \left(\frac{5.75}{10.25}, \frac{3.5}{10.25} \right).$$

Numerically,

$$\frac{5.75}{10.25} \approx 0.561, \quad \frac{3.5}{10.25} \approx 0.341.$$

So

$$\boxed{\hat{\mu} \approx (0.561, 0.341).}$$

(e) Predicted location after the next 2 seconds

Suppose the current location is approximately

$$(5.75, 3.5).$$

Random-walk prediction. If we use the random-walk drift from part (c), then one additional step would move us by about

$$\left(\frac{4}{5}, \frac{1}{5} \right).$$

So the predicted next location is

$$(5.75, 3.5) + \left(\frac{4}{5}, \frac{1}{5} \right) = (6.55, 3.70).$$

$$\boxed{(6.55, 3.70)}$$

Brownian-drift prediction over 2 seconds. If the Brownian drift is

$$\hat{\mu} \approx (0.561, 0.341)$$

per second, then over 2 seconds the expected displacement is

$$2\hat{\mu} \approx (1.122, 0.682).$$

So the predicted location is

$$(5.75, 3.5) + (1.122, 0.682) \approx (6.87, 4.18).$$

$$\boxed{(6.87, 4.18)}$$

Comment. The two predictions differ because the models are different:

- the random-walk model uses one discrete directional step,
- the Brownian model uses a continuous-time drift rate.

Problem 8

Let $(B_t)_{0 \leq t \leq 1}$ be standard Brownian motion.

We study the process conditioned on

$$B_1 = 0,$$

which is called a **Brownian bridge**.

(a) Joint distribution of (B_t, B_1)

Brownian motion is a Gaussian process. That means any finite collection of times has a multivariate normal distribution.

So (B_t, B_1) is bivariate normal.

Step 1: Compute the mean vector. Brownian motion has mean zero:

$$E[B_t] = 0, \quad E[B_1] = 0.$$

So the mean vector is

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Step 2: Compute the covariance matrix. For Brownian motion,

$$\text{Cov}(B_s, B_t) = \min(s, t).$$

Therefore

$$\text{Var}(B_t) = t, \quad \text{Var}(B_1) = 1, \quad \text{Cov}(B_t, B_1) = \min(t, 1) = t.$$

So

$$(B_t, B_1) \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} t & t \\ t & 1 \end{bmatrix} \right).$$

(b) Distribution of $B_t \mid B_1 = 0$

We use the conditional distribution formula for a multivariate normal vector.

Step 1: Conditional mean. For a bivariate normal pair (X, Y) ,

$$E[X \mid Y = y] = E[X] + \frac{\text{Cov}(X, Y)}{\text{Var}(Y)}(y - E[Y]).$$

Here

$$X = B_t, \quad Y = B_1, \quad y = 0.$$

So

$$E[B_t \mid B_1 = 0] = 0 + \frac{t}{1}(0 - 0) = 0.$$

Step 2: Conditional variance. For a bivariate normal pair,

$$\text{Var}(X | Y) = \text{Var}(X) - \frac{\text{Cov}(X, Y)^2}{\text{Var}(Y)}.$$

Thus

$$\text{Var}(B_t | B_1 = 0) = t - \frac{t^2}{1} = t - t^2.$$

So

$$\boxed{B_t | (B_1 = 0) \sim N(0, t - t^2).}$$

Step 3: Show an equivalent representation. Consider the process

$$B_t - tB_1.$$

Its mean is

$$E[B_t - tB_1] = 0.$$

Its variance is

$$\text{Var}(B_t - tB_1) = \text{Var}(B_t) + t^2 \text{Var}(B_1) - 2t \text{Cov}(B_t, B_1).$$

Substitute:

$$= t + t^2 - 2t^2 = t - t^2.$$

So $B_t - tB_1$ has the same one-dimensional distribution as $B_t | (B_1 = 0)$. In fact, one can show their finite-dimensional distributions agree, so

$$\boxed{(B_t | B_1 = 0)_{0 \leq t \leq 1} \stackrel{d}{=} (B_t - tB_1)_{0 \leq t \leq 1}.}$$

(c) Mean and covariance of the Brownian bridge

Using the representation

$$X_t = B_t - tB_1,$$

we can compute the mean and covariance directly.

Mean.

$$E[X_t] = E[B_t - tB_1] = 0.$$

So the mean function is

$$\boxed{m(t) = 0.}$$

Covariance. For $0 \leq s, t \leq 1$,

$$\text{Cov}(X_s, X_t) = \text{Cov}(B_s - sB_1, B_t - tB_1).$$

Expand this:

$$= \text{Cov}(B_s, B_t) - t \text{Cov}(B_s, B_1) - s \text{Cov}(B_1, B_t) + st \text{Var}(B_1).$$

Using Brownian covariance,

$$\text{Cov}(B_s, B_t) = \min(s, t), \quad \text{Cov}(B_s, B_1) = s, \quad \text{Cov}(B_t, B_1) = t, \quad \text{Var}(B_1) = 1.$$

So

$$\text{Cov}(X_s, X_t) = \min(s, t) - ts - st + st = \min(s, t) - st.$$

Thus the Brownian bridge has covariance function

$$\boxed{K(s, t) = \min(s, t) - st.}$$

Final answer. The Brownian bridge is a Gaussian process with

$$\boxed{m(t) = 0, \quad K(s, t) = \min(s, t) - st.}$$

(d) Drifted Brownian motion conditioned on $X_1 = 0$

Now let

$$X_t = \mu t + \sigma B_t.$$

We want the distribution of the process conditioned on

$$X_1 = 0.$$

Step 1: Use the condition $X_1 = 0$. Since

$$X_1 = \mu + \sigma B_1,$$

the condition $X_1 = 0$ implies

$$B_1 = -\frac{\mu}{\sigma}.$$

Now write

$$X_t = \mu t + \sigma B_t.$$

Add and subtract $\sigma t B_1$:

$$X_t = \mu t + \sigma(B_t - tB_1) + \sigma t B_1.$$

But under the conditioning,

$$\sigma t B_1 = \sigma t \left(-\frac{\mu}{\sigma}\right) = -\mu t.$$

So the drift terms cancel:

$$X_t \mid (X_1 = 0) = \sigma(B_t - tB_1).$$

Therefore

$$\boxed{X_t \mid (X_1 = 0) \stackrel{d}{=} \sigma(B_t - tB_1).}$$

Step 2: Compute mean and covariance. Since $B_t - tB_1$ has mean 0,

$$E[X_t \mid X_1 = 0] = 0.$$

Its covariance is just scaled by σ^2 :

$$\text{Cov}(X_s, X_t \mid X_1 = 0) = \sigma^2(\min(s, t) - st).$$

So

$$\boxed{E[X_t \mid X_1 = 0] = 0, \quad \text{Cov}(X_s, X_t \mid X_1 = 0) = \sigma^2(\min(s, t) - st).}$$

Interpretation. Conditioning a drifted Brownian motion to end at 0 at time 1 removes the deterministic drift in the conditional mean and leaves a scaled Brownian bridge.