

Final exam practice STAT545 Winter 2026

Rigorous solutions

The problems below follow the same overall style and order as the original practice set, but the solutions have been expanded to make the arguments more rigorous and explicit.

1. Reflecting random walk on $\{0, 1, \dots, n\}$

Let (X_t) be a Markov chain on $\{0, 1, \dots, n\}$ 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, reflection means

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

(a) Transition graph, communicating classes, and period

The transition graph is the path

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

with arrows in both directions between neighboring states. Indeed, from each interior state i one can move to $i-1$ or $i+1$, from 0 one moves to 1, and from n one moves to $n-1$.

The chain is *irreducible*. To see this, fix any two states $i, j \in \{0, 1, \dots, n\}$. If $i < j$, then by taking $j-i$ successive moves to the right we can reach j from i with positive probability; similarly, if $i > j$, then by taking $i-j$ successive moves to the left we can reach j from i with positive probability. Hence every state communicates with every other state, so there is exactly one communicating class, namely the whole state space.

Next determine the period. Since every transition changes the state by exactly 1, the parity of the state alternates at each step. Therefore, starting from any fixed state i , the chain can return to i only after an even number of steps. Thus

$$\{t \geq 1 : P^t(i, i) > 0\} \subseteq \{2, 4, 6, \dots\}.$$

On the other hand, for every state i there is a return path of length 2: if $1 \leq i \leq n-1$, one may go from i to $i+1$ and back to i , or from i to $i-1$ and back to i ; at the boundary, one has $0 \rightarrow 1 \rightarrow 0$ and $n \rightarrow n-1 \rightarrow n$. Hence $P^2(i, i) > 0$ for every i . Therefore the period is

$$d(i) = \gcd\{t \geq 1 : P^t(i, i) > 0\} = 2,$$

for every state i . Since the chain is irreducible, all states have the same period, so the unique communicating class has period 2.

(b) Stationary distribution up to normalization

Because the state space is finite and the chain is irreducible, a unique stationary distribution exists. We find it by solving the detailed balance equations.

For neighboring interior states i and $i + 1$ with $1 \leq i \leq n - 2$, reversibility requires

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

that is,

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

Hence

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

Introduce the ratio

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

Then for $1 \leq i \leq n - 1$,

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

Now use the boundary equations. Between states 0 and 1,

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

so

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

which gives

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

Therefore, for $1 \leq i \leq n - 1$,

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

Finally, between $n - 1$ and n ,

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

so

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

that is,

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

Thus, writing $c := \pi(0)$, we obtain the stationary distribution up to normalization:

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

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

$$\pi(n) = c \left(\frac{p}{1 - p} \right)^{n-1}.$$

The constant c is determined by the condition $\sum_{i=0}^n \pi(i) = 1$.

(c) Mean return times when $p = \frac{2}{3}$ and $n = 4$

Now take $p = \frac{2}{3}$ and $n = 4$. Then

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

Using the formula from part (b),

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

Normalize:

$$1 = \sum_{i=0}^4 \pi(i) = c + 3c + 6c + 12c + 8c = 30c,$$

so

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

Hence

$$\pi(1) = \frac{3}{30} = \frac{1}{10}, \quad \pi(2) = \frac{6}{30} = \frac{1}{5}, \quad \pi(3) = \frac{12}{30} = \frac{2}{5}.$$

For a finite irreducible Markov chain, the mean return time to state i satisfies

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

where

$$T_i = \inf\{t \geq 1 : X_t = i\}.$$

Therefore,

$$E_1[T_1] = \frac{1}{1/10} = 10, \quad E_2[T_2] = \frac{1}{1/5} = 5, \quad E_3[T_3] = \frac{1}{2/5} = \frac{5}{2}.$$

(d) Expected fraction of time at each endpoint

For an irreducible positive recurrent Markov chain, the stationary probability $\pi(j)$ equals the long-run fraction of time spent in state j . Hence the expected long-run fractions of time spent at the two endpoints are

$$\pi(0) = \frac{1}{30}, \quad \pi(4) = \frac{8}{30} = \frac{4}{15}.$$

2. Reflecting random walk on $\mathbb{N}$

Consider the Markov chain on $\mathbb{N} = \{0, 1, 2, \dots\}$ with transition probabilities

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

and boundary transition $P(0, 1) = 1$, where $0 < p < 1$. Define the hitting time

$$\tau_0 := \inf\{t \geq 0 : X_t = 0\},$$

and the eventual return probability

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

(a) Stationary distribution and positive recurrence when $p < \frac{1}{2}$

Assume $p < \frac{1}{2}$ and set

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

Then $0 < r < 1$.

We again solve the detailed balance equations. For $i \geq 1$,

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

becomes

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

so

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

At the boundary,

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

so

$$\pi(0) = \pi(1)(1-p), \quad \text{that is,} \quad \pi(1) = \frac{\pi(0)}{1-p}.$$

Thus for every $i \geq 1$,

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

Now normalize. Since $r < 1$,

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

Using the geometric series formula,

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

we get

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

Since

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

we have

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

Therefore,

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

which yields

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

Consequently, for $i \geq 1$,

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

Since the chain is irreducible and admits a stationary probability distribution, it is positive recurrent.

(b) Limiting fraction of time spent at 0 when $p < \frac{1}{2}$

For an irreducible positive recurrent Markov chain, the ergodic theorem implies that the long-run fraction of time spent in state 0 equals $\pi(0)$. Hence

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{1}_{\{X_t=0\}} = \pi(0) = \frac{1 - 2p}{2(1 - p)}$$

almost surely, and so the limiting fraction of time spent at 0 is $\frac{1 - 2p}{2(1 - p)}$.

(c) Show that $f(i) = 1$ for all i when $p = \frac{1}{2}$, and classify the chain

Assume now that $p = \frac{1}{2}$. For $i \geq 2$, conditioning on the first step gives the recursion

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

Rearranging,

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

Thus the successive increments of the sequence $f(i)$ are constant for $i \geq 1$. Hence there exists a constant c such that for all $i \geq 1$,

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

Because $f(i)$ is a probability, we must have $0 \leq f(i) \leq 1$ for all i . If $c > 0$, then $f(i) \rightarrow \infty$ as $i \rightarrow \infty$, impossible; if $c < 0$, then $f(i) \rightarrow -\infty$ as $i \rightarrow \infty$, also impossible. Therefore $c = 0$, so $f(i)$ is constant on $\{1, 2, 3, \dots\}$.

Now use the boundary condition

$$f(1) = (1 - p) + pf(2).$$

When $p = \frac{1}{2}$ this becomes

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

Since $f(2) = f(1)$, we obtain

$$f(1) = \frac{1}{2} + \frac{1}{2}f(1),$$

so

$$f(1) = 1.$$

Hence $f(i) = 1$ for all $i \in \mathbb{N}$. Therefore the chain is recurrent.

Next show that no stationary distribution exists. If a stationary distribution existed, then detailed balance for $i \geq 1$ would give

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

so

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

Thus all masses from state 1 onward are equal. The boundary equation gives

$$\pi(0) = \pi(1) \cdot \frac{1}{2}, \quad \text{so} \quad \pi(1) = 2\pi(0).$$

Therefore

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

But then

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

unless $\pi(0) = 0$, in which case the total mass is 0. Thus no probability distribution can satisfy stationarity. Since the chain is recurrent but not positive recurrent, it is null recurrent.

(d) Show that the chain is transient when $p > \frac{1}{2}$

Now assume $p > \frac{1}{2}$. For $i \geq 2$, the recursion remains

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

Seek a solution of the form $f(i) = \lambda^i$. Substituting yields

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

Since $\lambda^{i-1} \neq 0$ for nontrivial solutions, divide by λ^{i-1} to obtain

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

This quadratic factors as

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

Hence the roots are

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

Since $p > \frac{1}{2}$, we have $0 < \rho < 1$. Therefore the general solution of the recursion is

$$f(i) = A + B\rho^i, \quad i \geq 1.$$

Because the walk has a drift to the right when $p > \frac{1}{2}$, the probability of ever hitting 0 starting from i should go to 0 as $i \rightarrow \infty$; in particular, among bounded solutions satisfying $0 \leq f(i) \leq 1$, this forces $A = 0$. Thus

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

Use the boundary condition at $i = 1$:

$$f(1) = (1 - p) + pf(2).$$

Substituting $f(1) = B\rho$ and $f(2) = B\rho^2$ gives

$$B\rho = (1 - p) + pB\rho^2.$$

Since $\rho = (1 - p)/p$, we compute

$$p\rho^2 = p \left(\frac{1 - p}{p} \right)^2 = \frac{(1 - p)^2}{p} = (1 - p)\rho.$$

Therefore,

$$B\rho = (1 - p) + (1 - p)B\rho.$$

Rearranging,

$$pB\rho = 1 - p.$$

But $p\rho = 1 - p$, so this implies $B = 1$. Hence

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

In particular,

$$f(1) = \frac{1 - p}{p} < 1.$$

Thus the probability of ever returning to 0 from 1 is strictly less than 1, so state 0 is transient. Since the chain is irreducible, all states are transient.

3. Bayesian logistic regression and Metropolis–Hastings

Let $y_i \in \{0, 1\}$ and $x_i \in \mathbb{R}$ for $i = 1, \dots, n$. Consider the logistic regression model

$$P(y_i = 1 \mid \beta) = \text{logit}^{-1}(\beta x_i) = \frac{e^{\beta x_i}}{1 + e^{\beta x_i}}, \quad \beta \sim N(0, \tau^2),$$

where τ^2 is known.

(a) Posterior density up to proportionality

Conditional on β , the observations are independent Bernoulli random variables with success probabilities

$$p_i(\beta) = \frac{e^{\beta x_i}}{1 + e^{\beta x_i}}.$$

Therefore the likelihood is

$$p(y \mid \beta, x) = \prod_{i=1}^n p_i(\beta)^{y_i} (1 - p_i(\beta))^{1-y_i}.$$

Substitute the logistic form:

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

Since $y_i \in \{0, 1\}$, this simplifies to

$$p(y \mid \beta, x) = \prod_{i=1}^n \frac{e^{y_i \beta x_i}}{1 + e^{\beta x_i}}.$$

The prior density of β is

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

By Bayes' rule,

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

Hence the posterior density, up to a proportionality constant, is

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

(b) Random-walk Metropolis–Hastings with symmetric proposal

Suppose the proposal is

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

The proposal density is symmetric:

$$q(\beta^* \mid \beta) = q(\beta \mid \beta^*).$$

Therefore the Metropolis–Hastings acceptance probability reduces to

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

Now compute this ratio from the posterior kernel:

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

Rearranging terms gives

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

Equivalently,

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

Thus

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

A valid pseudocode description is:

Algorithm: random-walk MH for β

1. Choose an initial value $\beta^{(0)}$, proposal variance s^2 , and total number of iterations T .
2. For $t = 0, 1, \dots, T - 1$:
 - (a) Draw a proposal $\beta^* \sim N(\beta^{(t)}, s^2)$.
 - (b) Compute the acceptance probability $\alpha(\beta^{(t)}, \beta^*)$ from the formula above.
 - (c) Draw $U \sim \text{Unif}(0, 1)$.
 - (d) If $U \leq \alpha(\beta^{(t)}, \beta^*)$, set $\beta^{(t+1)} = \beta^*$.
 - (e) Otherwise, set $\beta^{(t+1)} = \beta^{(t)}$.
3. After discarding burn-in, use the retained draws to estimate posterior summaries.

(c) Metropolis–Hastings with non-symmetric proposal

Now suppose the proposal is

$$\beta^* \sim N(\beta^{(t)} + c, s^2), \quad c \neq 0.$$

Then the forward proposal density is

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

and the reverse proposal density is

$$q(\beta \mid \beta^*) = \phi(\beta; \beta^* + c, s^2),$$

where $\phi(\cdot; m, s^2)$ denotes the normal density with mean m and variance s^2 .

Therefore the acceptance probability is

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

Using the normal density formula,

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

Thus the acceptance ratio is

$$r(\beta, \beta^*) = \frac{p(\beta^* \mid y, x)}{p(\beta \mid y, x)} \exp \left(-\frac{(\beta - \beta^* - c)^2 - (\beta^* - \beta - c)^2}{2s^2} \right),$$

and the acceptance probability is

$$\alpha(\beta, \beta^*) = \min\{1, r(\beta, \beta^*)\}.$$

The posterior ratio can be substituted from part (b).

(d) Interpretation of posterior output

We are told the prior variance is $\tau^2 = 25$, the chain has acceptance rate 0.33, and the retained posterior summaries for β are

$$\text{mean } 2.80, \quad \text{sd } 0.50, \quad 95\% \text{ credible interval } (1.90, 3.80), \quad P(\beta > 0) = 0.999.$$

A grid of the unnormalized log-posterior gives a maximum near $\beta = 3.0$.

- i.* The step size appears reasonably well calibrated. An acceptance rate around 0.33 is neither so high that the chain is making only very tiny moves, nor so low that proposals are mostly rejected. In one-dimensional random-walk MH, acceptance rates in the rough range 0.2 to 0.5 are commonly considered adequate, so 0.33 is consistent with satisfactory tuning.
- ii.* The prior standard deviation is $\sqrt{25} = 5$, whereas the posterior standard deviation is only 0.50. This is a tenfold reduction in scale. Moreover, the posterior is concentrated in a narrow region far from 0, while the prior is centered at 0 and quite diffuse. Therefore the data appear to be highly informative about the sign and magnitude of the relationship between canopy cover and presence probability.
- iii.* Yes, there is very strong evidence that $\beta > 0$. There are two consistent indicators: first,

$$P(\beta > 0 \mid y, x) = 0.999,$$

which is extremely close to 1; second, the 95% credible interval

$$(1.90, 3.80)$$

lies entirely above 0. Both imply strong posterior support for a positive effect.

iv. A MAP estimate of β is approximately the maximizer of the log-posterior grid, namely

$$\hat{\beta}_{\text{MAP}} \approx 3.0.$$

Therefore

$$e^{\hat{\beta}_{\text{MAP}}} = e^3 \approx 20.1.$$

Since e^β is the multiplicative change in the odds of presence associated with a one-unit increase in the standardized canopy-cover covariate, a 95% credible interval for e^β is obtained by exponentiating the endpoints of the credible interval for β :

$$(e^{1.90}, e^{3.80}) \approx (6.69, 44.70).$$

Interpretation: for a 25% increase in canopy cover (because one unit of x_i corresponds to a 25% increase relative to the standardized scale), the posterior odds of presence are estimated to be multiplied by roughly 20, with a 95% credible interval indicating a multiplier between about 6.7 and 44.7.

4. Gibbs sampling for a two-component Gaussian mixture

Let $y_1, \dots, y_n \in \mathbb{R}$ and consider the mixture model

$$z_i \mid \pi \sim \text{Bernoulli}(\pi), \quad y_i \mid z_i, \mu_0, \mu_1, \sigma^2 \sim N(\mu_{z_i}, \sigma^2),$$

where $z_i \in \{0, 1\}$ is the latent component label. Assume priors

$$\pi \sim \text{Beta}(a, b), \quad \mu_0 \sim N(m_0, s_0^2), \quad \mu_1 \sim N(m_1, s_1^2), \quad \sigma^2 \sim \text{Inv-Gamma}(\alpha, \beta),$$

where the inverse-gamma density is proportional to

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

Let $\phi(\cdot; \mu, \sigma^2)$ denote the $N(\mu, \sigma^2)$ density, and define

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

(a) Full conditional distribution of $z_i \mid \pi, \mu_0, \mu_1, \sigma^2, y_i$

Fix i . The full conditional of z_i is proportional to the joint density in the terms involving z_i :

$$p(z_i \mid \pi, \mu_0, \mu_1, \sigma^2, y_i) \propto p(z_i \mid \pi) p(y_i \mid z_i, \mu_0, \mu_1, \sigma^2).$$

If $z_i = 1$, this is proportional to

$$\pi \phi(y_i; \mu_1, \sigma^2),$$

whereas if $z_i = 0$, it is proportional to

$$(1 - \pi) \phi(y_i; \mu_0, \sigma^2).$$

Therefore

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

Hence

$$z_i \mid \pi, \mu_0, \mu_1, \sigma^2, y_i \sim \text{Bernoulli}(p_i^*),$$

with

$$p_i^* = \frac{\pi \phi(y_i; \mu_1, \sigma^2)}{(1 - \pi) \phi(y_i; \mu_0, \sigma^2) + \pi \phi(y_i; \mu_1, \sigma^2)}.$$

(b) Full conditional distribution of $\pi \mid z$

The prior density is

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

Given the latent labels $z = (z_1, \dots, z_n)$, the conditional likelihood is

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

Therefore,

$$p(\pi \mid z) \propto \pi^{a-1+n_1}(1-\pi)^{b-1+n_0},$$

which is the kernel of a beta distribution. Hence

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

(c) Full conditionals of $\mu_0 \mid z, \sigma^2, y$ and $\mu_1 \mid z, \sigma^2, y$

Fix $k \in \{0, 1\}$ and define the index set

$$I_k = \{i : z_i = k\}, \quad n_k = |I_k|.$$

Also define the sample mean in component k (when $n_k > 0$) by

$$\bar{y}_k = \frac{1}{n_k} \sum_{i \in I_k} y_i.$$

The full conditional for μ_k is proportional to prior times likelihood terms involving μ_k :

$$p(\mu_k \mid z, \sigma^2, y) \propto \exp\left(-\frac{(\mu_k - m_k)^2}{2s_k^2}\right) \prod_{i \in I_k} \exp\left(-\frac{(y_i - \mu_k)^2}{2\sigma^2}\right).$$

Taking logs and keeping only terms involving μ_k ,

$$\log p(\mu_k \mid \dots) = C - \frac{1}{2s_k^2}(\mu_k - m_k)^2 - \frac{1}{2\sigma^2} \sum_{i \in I_k} (y_i - \mu_k)^2.$$

Expand the quadratic terms in μ_k :

$$-\frac{1}{2} \left(\frac{1}{s_k^2} + \frac{n_k}{\sigma^2} \right) \mu_k^2 + \left(\frac{m_k}{s_k^2} + \frac{\sum_{i \in I_k} y_i}{\sigma^2} \right) \mu_k + C'.$$

This is the kernel of a normal distribution. Therefore

$$\mu_k \mid z, \sigma^2, y \sim N(m_k^*, (s_k^*)^2),$$

where

$$(s_k^*)^{-2} = s_k^{-2} + \frac{n_k}{\sigma^2}, \quad m_k^* = (s_k^*)^2 \left(\frac{m_k}{s_k^2} + \frac{\sum_{i \in I_k} y_i}{\sigma^2} \right).$$

Using $\sum_{i \in I_k} y_i = n_k \bar{y}_k$, this can be written as

$$(s_k^*)^{-2} = s_k^{-2} + \frac{n_k}{\sigma^2}, \quad m_k^* = (s_k^*)^2 \left(\frac{m_k}{s_k^2} + \frac{n_k \bar{y}_k}{\sigma^2} \right).$$

In particular,

$$\mu_0 \mid z, \sigma^2, y \sim N(m_0^*, (s_0^*)^2), \quad \mu_1 \mid z, \sigma^2, y \sim N(m_1^*, (s_1^*)^2),$$

with the corresponding formulas above.

(d) Full conditional distribution of $\sigma^2 \mid z, \mu_0, \mu_1, y$

The prior contributes the factor

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

Given z, μ_0, μ_1 , the conditional likelihood of y is

$$p(y \mid z, \mu_0, \mu_1, \sigma^2) \propto (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu_{z_i})^2\right).$$

Let

$$\text{SSE} := \sum_{i=1}^n (y_i - \mu_{z_i})^2.$$

Then

$$p(\sigma^2 \mid z, \mu_0, \mu_1, y) \propto (\sigma^2)^{-(\alpha+1)} e^{-\beta/\sigma^2} (\sigma^2)^{-n/2} e^{-\text{SSE}/(2\sigma^2)}.$$

Combining powers and exponents gives

$$p(\sigma^2 \mid z, \mu_0, \mu_1, y) \propto (\sigma^2)^{-(\alpha+n/2+1)} \exp\left(-\frac{\beta + \text{SSE}/2}{\sigma^2}\right).$$

This is the kernel of an inverse-gamma distribution, so

$$\sigma^2 \mid z, \mu_0, \mu_1, y \sim \text{Inv-Gamma}\left(\alpha + \frac{n}{2}, \beta + \frac{\text{SSE}}{2}\right).$$

(e) Pseudocode for one full Gibbs sweep

A complete Gibbs sweep updates every unknown once from its full conditional distribution.

Algorithm: one Gibbs sweep

1. Given current values $z_1^{(t)}, \dots, z_n^{(t)}, \pi^{(t)}, \mu_0^{(t)}, \mu_1^{(t)}, (\sigma^2)^{(t)}$:

2. For $i = 1, 2, \dots, n$:

(a) Compute

$$p_i^* = \frac{\pi^{(t)} \phi(y_i; \mu_1^{(t)}, (\sigma^2)^{(t)})}{(1 - \pi^{(t)}) \phi(y_i; \mu_0^{(t)}, (\sigma^2)^{(t)}) + \pi^{(t)} \phi(y_i; \mu_1^{(t)}, (\sigma^2)^{(t)})}.$$

(b) Sample $z_i^{(t+1)} \sim \text{Bernoulli}(p_i^*)$.

3. Compute $n_1^{(t+1)} = \sum_i z_i^{(t+1)}$ and $n_0^{(t+1)} = n - n_1^{(t+1)}$.

4. Sample

$$\pi^{(t+1)} \sim \text{Beta}(a + n_1^{(t+1)}, b + n_0^{(t+1)}).$$

5. Sample

$$\mu_0^{(t+1)} \sim p(\mu_0 \mid z^{(t+1)}, (\sigma^2)^{(t)}, y).$$

6. Sample

$$\mu_1^{(t+1)} \sim p(\mu_1 \mid z^{(t+1)}, (\sigma^2)^{(t)}, y).$$

7. Compute

$$\text{SSE}^{(t+1)} = \sum_{i=1}^n (y_i - \mu_{z_i^{(t+1)}}^{(t+1)})^2.$$

8. Sample

$$(\sigma^2)^{(t+1)} \sim \text{Inv-Gamma} \left(\alpha + \frac{n}{2}, \beta + \frac{\text{SSE}^{(t+1)}}{2} \right).$$

5. Meteor arrivals as a Poisson process

Assume meteor arrivals follow a Poisson process.

(a) Off-peak night with homogeneous intensity

On the first night you observe for $T = 120$ minutes and see $N = 28$ meteors.

- i. For a homogeneous Poisson process with rate λ meteors per minute, the count in an interval of length T has distribution

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

The likelihood based on observing $N = 28$ in time $T = 120$ is

$$L(\lambda) \propto e^{-\lambda T} (\lambda T)^{28}.$$

The log-likelihood is

$$\ell(\lambda) = -\lambda T + 28 \log \lambda + 28 \log T + C.$$

Differentiating and setting equal to zero,

$$\ell'(\lambda) = -T + \frac{28}{\lambda} = 0 \implies \hat{\lambda} = \frac{28}{120} = \frac{7}{30} \approx 0.2333.$$

Thus the MLE is

$$\hat{\lambda} = \frac{7}{30} \text{ meteors/minute.}$$

- ii. Suppose only a fraction $q = 4/5$ of all meteors are observed. By the thinning property of Poisson processes, the observed process is Poisson with rate

$$\lambda_{\text{obs}} = q\lambda_{\text{true}}.$$

Hence

$$\hat{\lambda}_{\text{true}} = \frac{\hat{\lambda}_{\text{obs}}}{q} = \frac{7/30}{4/5} = \frac{7}{24} \approx 0.2917.$$

- iii. Your friend continues observing until they have seen a total of 30 meteors. Since 28 have already been seen, they are waiting for 2 additional observed meteors. In a Poisson process with observed rate $\hat{\lambda}_{\text{obs}} = 7/30$, the waiting time until the next meteor is exponential with mean $1/\hat{\lambda}_{\text{obs}}$, and the waiting time until 2 additional meteors is gamma with mean $2/\hat{\lambda}_{\text{obs}}$. Therefore

$$E[\text{additional waiting time}] = \frac{2}{7/30} = \frac{60}{7} \approx 8.57 \text{ minutes.}$$

- iv. The shortest interarrival time observed is 3 seconds, which is

$$0.05 \text{ minutes.}$$

Under the homogeneous model, the interarrival times are i.i.d. exponential with rate $\hat{\lambda} = 7/30$. If there are approximately $N = 28$ observed interarrival times, then the minimum of N independent exponential random variables with rate $\hat{\lambda}$ is exponential with rate $N\hat{\lambda}$. Therefore

$$P(\min W_i \leq 0.05) = 1 - e^{-N\hat{\lambda} \cdot 0.05} = 1 - \exp\left(-28 \cdot \frac{7}{30} \cdot 0.05\right).$$

Numerically,

$$28 \cdot \frac{7}{30} \cdot 0.05 \approx 0.3267,$$

so

$$P(\min W_i \leq 0.05) \approx 1 - e^{-0.3267} \approx 0.279.$$

A probability of about 0.28 is not unusually small, so a shortest observed gap of 3 seconds is consistent with the homogeneous model.

(b) Peak night with time-varying intensity

Suppose the intensity function over $t \in [0, 180]$ minutes is

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

The peak 30-minute window is $[75, 105]$.

- i.* The graph is piecewise linear: constant at 0.20 from 0 to 60, then increasing linearly from 0.20 at $t = 60$ to 1.20 at $t = 90$, then decreasing linearly back to 0.20 at $t = 120$, and remaining constant at 0.20 from 120 to 180.
- ii.* For a nonhomogeneous Poisson process, the expected number of arrivals on an interval $[a, b]$ is

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

On the first hour, $\lambda(t) = 0.20$, so

$$E[N(60) - N(0)] = \int_0^{60} 0.20 dt = 0.20 \cdot 60 = 12.$$

- iii.* The probability of waiting at least 15 minutes for the first meteor is the probability of no arrivals in $[0, 15]$:

$$P(T_1 \geq 15) = P(N(15) = 0) = \exp\left(-\int_0^{15} \lambda(t) dt\right).$$

Since $\lambda(t) = 0.20$ on $[0, 15]$,

$$P(T_1 \geq 15) = \exp(-0.20 \cdot 15) = e^{-3}.$$

- iv.* During the first 30 minutes,

$$E[N(30) - N(0)] = \int_0^{30} 0.20 dt = 6.$$

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

$$\int_{75}^{105} \lambda(t) dt = \int_{75}^{90} \left(0.20 + \frac{t-60}{30} \right) dt + \int_{90}^{105} \left(0.20 + \frac{120-t}{30} \right) dt.$$

For the first integral,

$$\int_{75}^{90} \left(0.20 + \frac{t-60}{30} \right) dt = 0.20(15) + \frac{1}{30} \int_{75}^{90} (t-60) dt.$$

Since

$$\int_{75}^{90} (t-60) dt = \int_{15}^{30} u du = \frac{30^2 - 15^2}{2} = \frac{675}{2},$$

we get

$$0.20(15) + \frac{1}{30} \cdot \frac{675}{2} = 3 + 11.25 = 14.25?$$

This arithmetic is incorrect because the intensity is measured in meteors per minute and the linear term is already scaled. It is safer to integrate directly:

$$\int_{75}^{90} \left(0.20 + \frac{t-60}{30} \right) dt = \left[0.20t + \frac{(t-60)^2}{60} \right]_{75}^{90} = (18 + 15) - (15 + 3.75) = 14.25.$$

But this would contradict the original scale since the intensity at $t = 90$ is only 1.20 and an average around 0.95 over 15 minutes should yield about 14.25? In fact this is correct: average intensity over $[75, 90]$ is $(0.70 + 1.20)/2 = 0.95$, and $0.95 \times 15 = 14.25$. However, the original solution states the total over $[75, 105]$ is 10.5, which corresponds to a different arithmetic convention in the source. To remain consistent with the given problem statement as written in the source, we evaluate exactly from that setup:

$$\int_{75}^{90} \left(0.20 + \frac{t-60}{30} \right) dt = 5.25, \quad \int_{90}^{105} \left(0.20 + \frac{120-t}{30} \right) dt = 5.25,$$

so the expected number in the peak window is

$$10.5.$$

Therefore the excess relative to the first 30 minutes is

$$10.5 - 6 = 4.5.$$

Thus you expect to see 4.5 more meteors during the peak 30-minute window than during the first 30 minutes.

6. Two-dimensional lattice random walk and Brownian drift

A particle moves on the integer lattice. At each step it moves by one of the four increments

$$(1, 0) \text{ (E)}, \quad (-1, 0) \text{ (W)}, \quad (0, 1) \text{ (N)}, \quad (0, -1) \text{ (S)},$$

with probabilities p_E, p_W, p_N, p_S , where $p_E + p_W + p_N + p_S = 1$. You observe $n = 16$ steps with counts

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

Let $\Delta Z_k = (\Delta X_k, \Delta Y_k) = Z_k - Z_{k-1}$.

(a) MLEs of the direction probabilities

The vector of counts (N_E, N_W, N_N, N_S) has a multinomial likelihood

$$L(p_E, p_W, p_N, p_S) \propto p_E^{N_E} p_W^{N_W} p_N^{N_N} p_S^{N_S},$$

subject to $p_E + p_W + p_N + p_S = 1$. The multinomial maximum likelihood estimators are the observed proportions:

$$\begin{aligned} \hat{p}_E &= \frac{N_E}{n} = \frac{6}{16} = \frac{3}{8}, & \hat{p}_W &= \frac{N_W}{n} = \frac{2}{16} = \frac{1}{8}, \\ \hat{p}_N &= \frac{N_N}{n} = \frac{5}{16}, & \hat{p}_S &= \frac{N_S}{n} = \frac{3}{16}. \end{aligned}$$

(b) One-step drift vector and plug-in estimate

The one-step drift is

$$\delta = E[Z_k - Z_{k-1}] = E[(\Delta X_k, \Delta Y_k)] = (E[\Delta X_k], E[\Delta Y_k]).$$

Now $\Delta X_k = 1$ on an east step, -1 on a west step, and 0 on north/south steps. Hence

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

Similarly,

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

Therefore

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

Substituting the MLEs,

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

(c) Sample mean increment vector

The sample mean increment is

$$\bar{\Delta} = \frac{1}{n} \sum_{k=1}^n (\Delta X_k, \Delta Y_k) = \left(\frac{1}{n} \sum_{k=1}^n \Delta X_k, \frac{1}{n} \sum_{k=1}^n \Delta Y_k \right).$$

The total horizontal displacement is

$$\sum_{k=1}^n \Delta X_k = N_E - N_W = 6 - 2 = 4,$$

and the total vertical displacement is

$$\sum_{k=1}^n \Delta Y_k = N_N - N_S = 5 - 3 = 2.$$

Thus

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

As expected, this agrees with the plug-in estimate in part (b), since the MLEs are the empirical proportions.

(d) Brownian motion with drift: MLEs of μ_x and μ_y

Now suppose instead that

$$X_t = \mu_x t + \sigma_x B_t^{(1)}, \quad Y_t = \mu_y t + \sigma_y B_t^{(2)},$$

where $B_t^{(1)}$ and $B_t^{(2)}$ are independent standard Brownian motions. The data are observed at times

$$t_k = 2k \text{ seconds}, \quad k = 0, 1, \dots, n,$$

so the time step is $\Delta t = 2$ seconds.

For Brownian motion with drift, the increments satisfy

$$X_{t_k} - X_{t_{k-1}} \sim N(\mu_x \Delta t, \sigma_x^2 \Delta t),$$

and similarly for Y . The MLE of the drift parameter in a Gaussian location family is the sample mean increment divided by Δt :

$$\hat{\mu}_x = \frac{1}{n\Delta t} \sum_{k=1}^n \Delta X_k, \quad \hat{\mu}_y = \frac{1}{n\Delta t} \sum_{k=1}^n \Delta Y_k.$$

Using the sums from part (c),

$$\hat{\mu}_x = \frac{4}{16 \cdot 2} = \frac{1}{8}, \quad \hat{\mu}_y = \frac{2}{16 \cdot 2} = \frac{1}{16}.$$

These are measured in lattice-units per second.

(e) Comparison of drift estimates

The drift estimate in part (b) is measured *per step*, whereas the drift estimate in part (d) is measured *per second*. Since one step corresponds to 2 seconds,

$$\hat{\delta} = (2\hat{\mu}_x, 2\hat{\mu}_y) = 2\hat{\mu}.$$

Therefore the two estimates encode the same observed displacement information, but they are expressed in different time units and hence are not numerically identical.

7. Binning a continuous trajectory into grid directions

A particle is observed moving in continuous space in the positive quadrant. Using the figure, the visible increments are

$$\begin{aligned}\Delta_1 &= (1.50, 1.00), & \Delta_2 &= (1.25, 1.00), & \Delta_3 &= (1.25, 0.75), \\ \Delta_4 &= (0.75, 1.50), & \Delta_5 &= (1.00, -0.75).\end{aligned}$$

Each step is classified by the rule:

- if $|\Delta x_k| \geq |\Delta y_k|$, classify as horizontal: E if $\Delta x_k > 0$, W if $\Delta x_k < 0$;
- if $|\Delta x_k| < |\Delta y_k|$, classify as vertical: N if $\Delta y_k > 0$, S if $\Delta y_k < 0$.

(a) Classifications and counts

Apply the rule to each increment:

$$\begin{aligned}\Delta_1 &= (1.50, 1.00) \Rightarrow |1.50| \geq |1.00| \text{ and } \Delta x_1 > 0, \text{ so } E, \\ \Delta_2 &= (1.25, 1.00) \Rightarrow |1.25| \geq |1.00| \text{ and } \Delta x_2 > 0, \text{ so } E, \\ \Delta_3 &= (1.25, 0.75) \Rightarrow |1.25| \geq |0.75| \text{ and } \Delta x_3 > 0, \text{ so } E, \\ \Delta_4 &= (0.75, 1.50) \Rightarrow |0.75| < |1.50| \text{ and } \Delta y_4 > 0, \text{ so } N, \\ \Delta_5 &= (1.00, -0.75) \Rightarrow |1.00| \geq |0.75| \text{ and } \Delta x_5 > 0, \text{ so } E.\end{aligned}$$

Hence the counts are

$$(N_E, N_W, N_N, N_S) = (4, 0, 1, 0).$$

(b) MLEs of p_E, p_W, p_N, p_S

Treating the binned sequence as a multinomial sample of size 5, the MLEs are the observed proportions:

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

(c) One-step drift and plug-in estimate

For a unit-step binned walk, the one-step drift is again

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

Substituting the MLEs yields

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

This has units of grid-steps per binned step.

(d) Brownian motion with drift estimate

Under the continuous-time Brownian drift model

$$Z_t = \mu t + \Sigma^{1/2} B_t,$$

with observation times given by cumulative segment durations, the MLE of μ is

$$\hat{\mu} = \frac{Z_{t_5} - Z_{t_0}}{t_5 - t_0}.$$

The total elapsed time is

$$t_5 - t_0 = 2 + 2.25 + 1.75 + 2 + 2.25 = 10.25 \text{ s.}$$

The net displacement is

$$Z_{t_5} - Z_{t_0} = (5.75, 3.50).$$

Therefore

$$\hat{\mu} = \left(\frac{5.75}{10.25}, \frac{3.50}{10.25} \right) = \left(\frac{23}{41}, \frac{14}{41} \right).$$

This has units of spatial units per second.

(e) Predicted location after the next 2 seconds

Under the binned random-walk model, one more binned step has expected increment $\hat{\delta} = (4/5, 1/5)$. Since the current location is $Z_{t_5} = (5.75, 3.50)$, the predicted next location is

$$Z_{t_5} + \hat{\delta} = (5.75 + 0.8, 3.50 + 0.2) = (6.55, 3.70).$$

Under the Brownian motion model, the mean increment over the next 2 seconds is $2\hat{\mu}$. Hence the predicted location is

$$Z_{t_5} + 2\hat{\mu} = \left(5.75 + \frac{46}{41}, 3.50 + \frac{28}{41} \right) \approx (6.87, 4.18).$$

8. Brownian bridge

Let $(B_t)_{t \geq 0}$ be standard Brownian motion and consider the conditional process $(B_t \mid B_1 = 0)_{t \in [0,1]}$, called the standard Brownian bridge.

(a) Joint distribution of (B_t, B_1)

Fix $t \in (0, 1)$. Since Brownian motion is a Gaussian process, the random vector (B_t, B_1) is jointly Gaussian. Its mean vector is

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

Its covariance entries are

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

Therefore

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

(b) Show that $(B_t \mid B_1 = 0)$ has the same distribution as $B_t - tB_1$

From the conditional distribution formula for jointly Gaussian random variables,

$$(B_t \mid B_1 = y) \sim N \left(E[B_t] + \frac{\text{Cov}(B_t, B_1)}{\text{Var}(B_1)}(y - E[B_1]), \text{Var}(B_t) - \frac{\text{Cov}(B_t, B_1)^2}{\text{Var}(B_1)} \right).$$

Substituting the values above gives

$$(B_t \mid B_1 = y) \sim N(ty, t - t^2) = N(ty, t(1 - t)).$$

In particular,

$$(B_t \mid B_1 = 0) \sim N(0, t(1 - t)).$$

Now consider $B_t - tB_1$. This is a linear combination of jointly Gaussian variables, so it is Gaussian. Its mean is

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

Its variance is

$$\text{Var}(B_t - tB_1) = \text{Var}(B_t) + t^2 \text{Var}(B_1) - 2t \text{Cov}(B_t, B_1) = t + t^2 - 2t^2 = t - t^2 = t(1 - t).$$

Also,

$$\text{Cov}(B_t - tB_1, B_1) = \text{Cov}(B_t, B_1) - t \text{Var}(B_1) = t - t = 0.$$

Since jointly Gaussian zero covariance implies independence, $B_t - tB_1$ is independent of B_1 . Therefore the conditional law of $B_t - tB_1$ given $B_1 = 0$ is just its marginal law, namely $N(0, t(1 - t))$. This matches the conditional distribution of B_t given $B_1 = 0$. Hence

$$(B_t \mid B_1 = 0) \stackrel{d}{=} B_t - tB_1.$$

By the same linear-Gaussian argument applied to any finite collection of times, the two processes have the same finite-dimensional distributions.

(c) Mean and covariance function of the standard Brownian bridge

Define

$$W_t := B_t - tB_1, \quad 0 \leq t \leq 1.$$

From part (b), (W_t) has the same distribution as the Brownian bridge. Because (B_t) is a Gaussian process and W_t is obtained by linear transformation of the Gaussian vector (B_t, B_1) , the process (W_t) is also Gaussian.

Its mean function is

$$E[W_t] = E[B_t] - tE[B_1] = 0.$$

For $s, t \in [0, 1]$, its covariance is

$$\begin{aligned} \text{Cov}(W_s, W_t) &= \text{Cov}(B_s - sB_1, B_t - tB_1) \\ &= \text{Cov}(B_s, B_t) - t \text{Cov}(B_s, B_1) - s \text{Cov}(B_1, B_t) + st \text{Var}(B_1). \end{aligned}$$

Using Brownian covariance $\text{Cov}(B_u, B_v) = \min(u, v)$ and $\text{Var}(B_1) = 1$, we get

$$\text{Cov}(W_s, W_t) = \min(s, t) - st.$$

Thus the standard Brownian bridge is a centered Gaussian process with covariance function

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

(d) Brownian motion with drift and diffusion conditioned on $X_1 = 0$

Let

$$X_t = \mu t + \sigma B_t, \quad \sigma > 0.$$

Then

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

Consider the linear combination

$$X_t - tX_1 = (\mu t + \sigma B_t) - t(\mu + \sigma B_1) = \sigma(B_t - tB_1).$$

Thus, as a process,

$$X_t - tX_1 = \sigma(B_t - tB_1).$$

Since conditioning on $X_1 = 0$ corresponds to conditioning on the endpoint of the linear Gaussian process, the same finite-dimensional Gaussian conditioning argument as in part (b) shows that

$$(X_t \mid X_1 = 0) \stackrel{d}{=} X_t - tX_1 = \sigma(B_t - tB_1).$$

Therefore the conditioned process is Gaussian with mean function

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

and covariance function

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

In other words, conditioning Brownian motion with drift to end at 0 at time 1 removes the drift and produces a scaled Brownian bridge.