UCLA-ALIGNED COURSE FIELD NOTES

STATS 102C

Introduction to Monte Carlo Methods

Simulation, importance sampling, Markov chains, MCMC, diagnostics, and advanced Monte Carlo computation

How these notes are written. I wrote each chapter the way I wish the material had been explained the first time: the real idea, the point where I usually get stuck, the simplest mental model that still stays honest, and a worked decision instead of a polished answer appearing from nowhere. Every chapter closes with practice and a fully written solution. If a sentence sounds impressive but does not help me solve or explain something, it does not belong here.

Daniel Mastick

Curriculum map, working notes, practice, and solutions

Course map

Week	Core thread	What should be solid by the end
1	**Monte Carlo integration and error**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
2	**Random variate generation**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
3	**Variance reduction and importance sampling**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
4	**Markov chains and stationary distributions**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
5	**Metropolis-Hastings algorithms**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
6	**Gibbs sampling and data augmentation**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
7	**MCMC convergence and effective sample size**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
8	**Tempering, sequential ideas, and trustworthy workflows**	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
9	**Integration studio**	Combine several chapters in one case, preserve their assumptions, and reconcile the result across representations.
10	**Cumulative review**	Retrieve the full course map from memory and solve mixed problems without chapter labels.

Scope anchor: https://catalog.registrar.ucla.edu/course/2025/STATS102C. The sequence is aligned to the official catalog description and expanded into a practical ten-week study plan.

How I would use this packet

Treat each numbered section as one chapter. Read the formal explanation, pause at the student interpretation, and see whether the easy version still says the same thing. Rework the example without looking, then cover the solution in the chapter practice box. The cumulative set at the end is deliberately mixed: knowledge becomes durable when I have to choose the tool instead of being told which chapter I am in.

Contents

CHAPTER 1 Monte Carlo integration and error

The idea

Monte Carlo replaces an expectation or integral with an average of simulated values. Unbiasedness, variance, the central limit theorem, confidence intervals, and square-root convergence quantify simulation error independently of dimension.

Rebuild the chapter from first principles

The claim I need to own is this: Monte Carlo replaces an expectation or integral with an average of simulated values. Unbiasedness, variance, the central limit theorem, confidence intervals, and square-root convergence quantify simulation error independently of dimension. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of monte carlo integration and error. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach monte carlo integration and error on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from the course foundations to random variate generation. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Monte Carlo replaces an expectation or integral with an average of simulated values.
- **What moves.** Unbiasedness, variance, the central limit theorem, confidence intervals, and square-root convergence quantify simulation error independently of dimension.
- **What keeps the answer honest.** The model becomes useful only when I can apply monte carlo integration and error to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

I report Monte Carlo error beside the estimate because the simulation is another sampling procedure.

The easy version

Monte Carlo replaces an expectation or integral with an average of simulated values.

Worked example

Question. How many times more draws roughly halve ordinary Monte Carlo standard error?
Walkthrough. Four times as many, because standard error scales as one over square root of the number of independent draws.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. How many times more draws roughly halve ordinary Monte Carlo standard error?
Solution. Four times as many, because standard error scales as one over square root of the number of independent draws.

Practice 2. Explain monte carlo integration and error to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct monte carlo integration and error from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a monte carlo integration and error problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 2 Random variate generation

The idea

Inverse transforms, composition, acceptance-rejection, and transformation methods turn uniform pseudo-random numbers into target distributions. Correct support, envelope constants, acceptance rates, and diagnostic comparisons verify the generator.

Rebuild the chapter from first principles

The claim I need to own is this: Inverse transforms, composition, acceptance-rejection, and transformation methods turn uniform pseudo-random numbers into target distributions. Correct support, envelope constants, acceptance rates, and diagnostic comparisons verify the generator. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of random variate generation. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach random variate generation on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from monte carlo integration and error to variance reduction and importance sampling. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Inverse transforms, composition, acceptance-rejection, and transformation methods turn uniform pseudo-random numbers into target distributions.
- **What moves.** Correct support, envelope constants, acceptance rates, and diagnostic comparisons verify the generator.

- **What keeps the answer honest.** The model becomes useful only when I can apply random variate generation to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

I test a sampler on moments, quantiles, and plots before using it inside a larger algorithm.

The easy version

Inverse transforms, composition, acceptance-rejection, and transformation methods turn uniform pseudo-random numbers into target distributions.

Worked example

Question. In rejection sampling, why must M g(x) dominate f(x)?
Walkthrough. The acceptance probability f(x)/(M g(x)) must stay at most one everywhere. If domination fails, the algorithm is not sampling the intended target.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. In rejection sampling, why must M g(x) dominate f(x)?
Solution. The acceptance probability f(x)/(M g(x)) must stay at most one everywhere. If domination fails, the algorithm is not sampling the intended target.

Practice 2. Explain random variate generation to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct random variate generation from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a random variate generation problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output?

Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 3 Variance reduction and importance sampling

The idea

Antithetic variables, control variates, stratification, Rao-Blackwellization, and importance sampling reduce Monte Carlo variance by using structure. Importance weights require target support coverage and can collapse when proposal tails are too light.

Rebuild the chapter from first principles

The claim I need to own is this: Antithetic variables, control variates, stratification, Rao-Blackwellization, and importance sampling reduce Monte Carlo variance by using structure. Importance weights require target support coverage and can collapse when proposal tails are too light. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of variance reduction and importance sampling. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes,

the invariant breaks, or the model stops matching reality.

- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach variance reduction and importance sampling on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from random variate generation to markov chains and stationary distributions. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Antithetic variables, control variates, stratification, Rao-Blackwellization, and importance sampling reduce Monte Carlo variance by using structure.
- **What moves.** Importance weights require target support coverage and can collapse when proposal tails are too light.
- **What keeps the answer honest.** The model becomes useful only when I can apply variance reduction and importance sampling to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

My shorthand for this chapter is: the best simulation is often the one that wastes less randomness, not merely the one that runs longer.

The easy version

Antithetic variables, control variates, stratification, Rao-Blackwellization, and importance sampling reduce Monte Carlo variance by using structure.

Worked example

Question. What makes an effective control variate?
Walkthrough. It has known expectation and strong correlation with the target quantity. Subtracting an optimally scaled centered control removes shared variation without changing expectation.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. What makes an effective control variate?
Solution. It has known expectation and strong correlation with the target quantity. Subtracting an optimally scaled centered control removes shared variation without changing expectation.

Practice 2. Explain variance reduction and importance sampling to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct variance reduction and importance sampling from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a variance reduction and importance sampling problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 4 Markov chains and stationary distributions

The idea

A Markov chain evolves through a transition kernel depending only on current state. Irreducibility, aperiodicity, recurrence, detailed balance, and stationarity govern whether long-run averages represent a target distribution.

Rebuild the chapter from first principles

The claim I need to own is this: A Markov chain evolves through a transition kernel depending only on current state. Irreducibility, aperiodicity, recurrence, detailed balance, and stationarity govern whether long-run averages represent a target distribution. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of markov chains and stationary distributions. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach markov chains and stationary distributions on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from variance reduction and importance sampling to metropolis-hastings algorithms. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** A Markov chain evolves through a transition kernel depending only on current state.
- **What moves.** Irreducibility, aperiodicity, recurrence, detailed balance, and stationarity govern

whether long-run averages represent a target distribution.

- **What keeps the answer honest.** The model becomes useful only when I can apply markov chains and stationary distributions to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

MCMC draws are dependent on purpose, so I study the chain rather than pretending I received iid samples.

The easy version

A Markov chain evolves through a transition kernel depending only on current state.

Worked example

Question. What does detailed balance imply?
Walkthrough. If pi(x)P(x,y)=pi(y)P(y,x), then probability flow balances pairwise and pi is stationary, though convergence also needs appropriate chain structure.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. What does detailed balance imply?
Solution. If pi(x)P(x,y)=pi(y)P(y,x), then probability flow balances pairwise and pi is stationary, though convergence also needs appropriate chain structure.

Practice 2. Explain markov chains and stationary distributions to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct markov chains and stationary distributions from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a markov chains and stationary distributions problem but never states a model or checks the result. Write the shortest useful review of that solution.

Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 5 Metropolis-Hastings algorithms

The idea

Metropolis-Hastings proposes a move and accepts it using a ratio correcting target and proposal asymmetry. Random-walk scale controls exploration; log densities prevent underflow; initialization and multimodality affect convergence.

Rebuild the chapter from first principles

The claim I need to own is this: Metropolis-Hastings proposes a move and accepts it using a ratio correcting target and proposal asymmetry. Random-walk scale controls exploration; log densities prevent underflow; initialization and multimodality affect convergence. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of metropolis-hastings algorithms. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.

- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach metropolis-hastings algorithms on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from markov chains and stationary distributions to gibbs sampling and data augmentation. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Metropolis-Hastings proposes a move and accepts it using a ratio correcting target and proposal asymmetry.
- **What moves.** Random-walk scale controls exploration; log densities prevent underflow; initialization and multimodality affect convergence.
- **What keeps the answer honest.** The model becomes useful only when I can apply metropolis-hastings algorithms to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

A high acceptance rate can mean tiny useless steps, while a low rate can mean the chain barely moves. I judge exploration, not one percentage.

The easy version

Metropolis-Hastings proposes a move and accepts it using a ratio correcting target and proposal asymmetry.

Worked example

Question. Why may an unnormalized target density be used?
Walkthrough. The unknown normalizing constant appears in both numerator and denominator of the acceptance ratio and cancels.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. Why may an unnormalized target density be used?
Solution. The unknown normalizing constant appears in both numerator and denominator of the acceptance ratio and cancels.

Practice 2. Explain metropolis-hastings algorithms to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct metropolis-hastings algorithms from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a metropolis-hastings algorithms problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 6 Gibbs sampling and data augmentation

The idea

Gibbs sampling updates components from full conditional distributions and automatically accepts. Blocking, parameterization, conjugacy, latent-variable augmentation, and conditional dependence determine mixing speed.

Rebuild the chapter from first principles

The claim I need to own is this: Gibbs sampling updates components from full conditional distributions and automatically accepts. Blocking, parameterization, conjugacy, latent-variable augmentation, and conditional dependence determine mixing speed. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of gibbs sampling and data augmentation. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach gibbs sampling and data augmentation on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from metropolis-hastings algorithms to mcmc convergence and effective sample size. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Gibbs sampling updates components from full conditional distributions and automatically accepts.

- **What moves.** Blocking, parameterization, conjugacy, latent-variable augmentation, and conditional dependence determine mixing speed.
- **What keeps the answer honest.** The model becomes useful only when I can apply gibbs sampling and data augmentation to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

My shorthand for this chapter is: easy conditional draws do not guarantee an efficient chain; strongly coupled coordinates can still crawl.

The easy version

Gibbs sampling updates components from full conditional distributions and automatically accepts.

Worked example

Question. When can blocking improve Gibbs sampling?
Walkthrough. Jointly updating highly correlated parameters can move across their dependence structure more efficiently than alternating small conditional moves.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. When can blocking improve Gibbs sampling?
Solution. Jointly updating highly correlated parameters can move across their dependence structure more efficiently than alternating small conditional moves.

Practice 2. Explain gibbs sampling and data augmentation to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct gibbs sampling and data augmentation from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a gibbs sampling and data

augmentation problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 7 MCMC convergence and effective sample size

The idea

Trace plots, autocorrelation, multiple chains, rank diagnostics, effective sample size, Monte Carlo standard error, and posterior predictive checks diagnose computation. Burn-in alone cannot repair a poorly mixing or nonconverged chain.

Rebuild the chapter from first principles

The claim I need to own is this: Trace plots, autocorrelation, multiple chains, rank diagnostics, effective sample size, Monte Carlo standard error, and posterior predictive checks diagnose computation. Burn-in alone cannot repair a poorly mixing or nonconverged chain. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of mcmc convergence and effective sample size. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.

- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach mcmc convergence and effective sample size on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from gibbs sampling and data augmentation to tempering, sequential ideas, and trustworthy workflows. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Trace plots, autocorrelation, multiple chains, rank diagnostics, effective sample size, Monte Carlo standard error, and posterior predictive checks diagnose computation.
- **What moves.** Burn-in alone cannot repair a poorly mixing or nonconverged chain.
- **What keeps the answer honest.** The model becomes useful only when I can apply mcmc convergence and effective sample size to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

I do not call a chain converged because the trace looks fuzzy; I want independent chains to agree and enough effective information for the claim.

The easy version

Trace plots, autocorrelation, multiple chains, rank diagnostics, effective sample size, Monte Carlo standard error, and posterior predictive checks diagnose computation.

Worked example

Question. Why is 10,000 MCMC iterations not equivalent to 10,000 iid draws?
Walkthrough. Autocorrelation reduces independent information. Effective sample size can be far smaller, so uncertainty must account for chain dependence.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. Why is 10,000 MCMC iterations not equivalent to 10,000 iid draws?
Solution. Autocorrelation reduces independent information. Effective sample size can be far smaller, so uncertainty must account for chain dependence.

Practice 2. Explain mcmc convergence and effective sample size to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct mcmc convergence and effective sample size from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a mcmc convergence and effective sample size problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

CHAPTER 8 Tempering, sequential ideas, and trustworthy workflows

The idea

Simulated tempering and parallel tempering flatten targets to cross modes; sequential Monte Carlo moves weighted particles through distributions. Reproducible seeds, vectorized log probability, diagnostics, sensitivity, and benchmark cases make advanced samplers trustworthy.

Rebuild the chapter from first principles

The claim I need to own is this: Simulated tempering and parallel tempering flatten targets to cross modes; sequential Monte Carlo moves weighted particles through distributions. Reproducible seeds, vectorized log probability, diagnostics, sensitivity, and benchmark cases make advanced samplers trustworthy. I do not count that as learned just because the sentence looks familiar. I should be able to name the objects, state what changes and what stays fixed, recover the rule from the definitions, and explain why the result is allowed.

The move

My working sequence

1. **Translate.** Rewrite the prompt in the language of tempering, sequential ideas, and trustworthy workflows. List the known quantities, the unknown, the units or types, and any hidden constraint.
2. **Choose.** State the theorem, model, algorithm, accounting rule, or physical law before substituting values. This is where I make the assumption visible.
3. **Execute.** Work one justified step at a time and keep intermediate state readable enough that I can audit it later.
4. **Verify.** Check a boundary case, sign, unit, invariant, order of magnitude, or alternative derivation. Then translate the result back into an ordinary sentence.

Details I would refuse to skip

- **Definitions:** I should be able to define every technical noun in the chapter without using that same noun in the definition.
- **Assumptions:** I write the condition that makes the method legal beside the method itself. A correct procedure under the wrong assumptions is still a wrong answer.
- **Representation:** I move freely among words, equations or code, a diagram, and a small numerical example. If one representation disagrees with another, that disagreement is useful evidence.
- **Failure modes:** I keep the nearest counterexample in mind: the boundary where the rule changes, the invariant breaks, or the model stops matching reality.
- **Communication:** My final answer names what the result means and what it does not establish. I do not let notation do the explanatory work for me.

Sanity check

Closed-note check. Can I teach tempering, sequential ideas, and trustworthy workflows on a blank page, derive or reconstruct its main rule, work a fresh example, and diagnose a deliberately broken solution? If I can only recognize the finished work, I am not done.

How this chapter connects

I read this chapter as the bridge from mcmc convergence and effective sample size to the cumulative course model. The earlier material supplies the language and constraints; this chapter adds a new reasoning move; the later material asks me to use that move without being told its name. That connection matters because exams and real projects mix chapters on purpose.

The full working model

- **Core claim.** Simulated tempering and parallel tempering flatten targets to cross modes; sequential Monte Carlo moves weighted particles through distributions.
- **What moves.** Reproducible seeds, vectorized log probability, diagnostics, sensitivity, and benchmark cases make advanced samplers trustworthy.
- **What keeps the answer honest.** The model becomes useful only when I can apply tempering, sequential ideas, and trustworthy workflows to an unfamiliar case and defend the assumptions.
- **Mastery standard.** I can define the objects, rebuild the rule, work a fresh case, compare alternatives, and diagnose a broken solution without relying on recognition.

How I actually think about it

I first prove the code on a target whose answer I know, then earn the right to use it on the hard model.

The easy version

Simulated tempering and parallel tempering flatten targets to cross modes; sequential Monte Carlo moves weighted particles through distributions.

Worked example

Question. How does parallel tempering help a multimodal target?
Walkthrough. Hot chains explore a flatter landscape and cross energy barriers. Swaps can carry discoveries back to the cold target chain while preserving the joint stationary construction.

Common miss

The fastest way to get this chapter wrong is to use the visible procedure without naming its assumption, units, time period, reference group, or boundary. I put that condition beside the work and verify the result independently.

Solved chapter practice

Practice 1. How does parallel tempering help a multimodal target?
Solution. Hot chains explore a flatter landscape and cross energy barriers. Swaps can carry discoveries back to the cold target chain while preserving the joint stationary construction.

Practice 2. Explain tempering, sequential ideas, and trustworthy workflows to a classmate using one definition, one concrete example, one assumption, and one failure mode.
Solution. Begin with the core claim above, use the worked case as the concrete example, state the condition highlighted by the warning, and choose a failure where that condition does not hold. A complete explanation connects all four pieces instead of listing them.

Mastery practice A. Reconstruct tempering, sequential ideas, and trustworthy workflows from a blank page. Include the central definition, the governing rule, one required assumption, and one boundary case.
Solution. Start from the blue idea box, but rewrite it in my own words. Identify the objects

and state, name the rule before using it, copy the relevant assumption from the details audit, and choose a boundary where that assumption becomes equality or fails. A complete solution explains why each part belongs; a list of vocabulary is not enough.

Mastery practice B. A classmate gets a polished-looking answer to a tempering, sequential ideas, and trustworthy workflows problem but never states a model or checks the result. Write the shortest useful review of that solution.
Solution. I would ask four concrete questions: What exactly are the inputs and desired output? Which rule licenses the main step? Which assumption could invalidate it? Which independent check supports the conclusion? If the answer cannot survive those four questions, I treat the numerical or verbal result as unverified.

Cumulative study plan

I use this packet in three passes. First I reconstruct each model and work its examples without looking. Second I mix chapters so the tool is no longer named for me. Third I explain a complete case aloud, including assumptions, units, uncertainty, failure modes, and what evidence would change my conclusion.