

UCLA-ALIGNED COURSE FIELD NOTES

STATS 100A

Introduction to Probability

Counting, conditioning, random variables, joint laws, expectation, transformations, and limit theorems

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	Probability spaces and counting	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
2	Conditional probability, independence, and Bayes	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
3	Discrete random variables	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
4	Continuous random variables	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
5	Joint distributions and random vectors	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
6	Functions and transformations of variables	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
7	Expectation, conditioning, and generating tools	Reconstruct the model, work an application, test its assumptions, and explain the result in ordinary language.
8	Laws of large numbers and central limit theorem	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/STATS100A?year=2025>. 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

1	Probability spaces and counting	5
1.1	Rebuild the chapter from first principles	5
1.2	Details I would refuse to skip	5
1.3	How this chapter connects	6
1.4	The full working model	6
2	Conditional probability, independence, and Bayes	7
2.1	Rebuild the chapter from first principles	7
2.2	Details I would refuse to skip	8
2.3	How this chapter connects	8
2.4	The full working model	8
3	Discrete random variables	10
3.1	Rebuild the chapter from first principles	10
3.2	Details I would refuse to skip	10
3.3	How this chapter connects	11
3.4	The full working model	11
4	Continuous random variables	12
4.1	Rebuild the chapter from first principles	12
4.2	Details I would refuse to skip	13
4.3	How this chapter connects	13
4.4	The full working model	13
5	Joint distributions and random vectors	15
5.1	Rebuild the chapter from first principles	15
5.2	Details I would refuse to skip	15
5.3	How this chapter connects	16
5.4	The full working model	16
6	Functions and transformations of variables	17
6.1	Rebuild the chapter from first principles	17
6.2	Details I would refuse to skip	18
6.3	How this chapter connects	18

6.4	The full working model	18
7	Expectation, conditioning, and generating tools	20
7.1	Rebuild the chapter from first principles	20
7.2	Details I would refuse to skip	20
7.3	How this chapter connects	21
7.4	The full working model	21
8	Laws of large numbers and central limit theorem	22
8.1	Rebuild the chapter from first principles	22
8.2	Details I would refuse to skip	23
8.3	How this chapter connects	23
8.4	The full working model	23

CHAPTER 1 Probability spaces and counting

The idea

A probability model specifies outcomes, events, and a measure obeying nonnegativity, normalization, and additivity. Permutations, combinations, inclusion-exclusion, and symmetry calculate finite models only after the sample space is defined correctly.

Rebuild the chapter from first principles

The claim I need to own is this: A probability model specifies outcomes, events, and a measure obeying nonnegativity, normalization, and additivity. Permutations, combinations, inclusion-exclusion, and symmetry calculate finite models only after the sample space is defined correctly. 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 probability spaces and counting. 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 probability spaces and counting 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 conditional probability, independence, and bayes. 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 probability model specifies outcomes, events, and a measure obeying nonnegativity, normalization, and additivity.
- **What moves.** Permutations, combinations, inclusion-exclusion, and symmetry calculate finite models only after the sample space is defined correctly.
- **What keeps the answer honest.** The model becomes useful only when I can apply probability spaces and counting 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 write what one equally likely outcome is before counting anything.

The easy version

A probability model specifies outcomes, events, and a measure obeying nonnegativity, normalization, and additivity.

Worked example

Question. Five cards are drawn from 52 without order. How many hands exist?

Walkthrough. There are $52 \text{ choose } 5$, equal to 2,598,960. Combinations apply because order within a hand does not matter.

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. Five cards are drawn from 52 without order. How many hands exist?

Solution. There are 52 choose 5, equal to 2,598,960. Combinations apply because order within a hand does not matter.

Practice 2. Explain probability spaces and counting 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 probability spaces and counting 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 probability spaces and counting 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 Conditional probability, independence, and Bayes

The idea

Conditioning restricts the sample space using new information. Independence is the factorization of joint probability, not the absence of any relationship in ordinary language; Bayes' rule reverses conditioning using base rates.

Rebuild the chapter from first principles

The claim I need to own is this: Conditioning restricts the sample space using new information. Independence is the factorization of joint probability, not the absence of any relationship in ordinary language; Bayes' rule reverses conditioning using base rates. 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 conditional probability, independence, and bayes. 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 conditional probability, independence, and bayes 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 probability spaces and counting to discrete random variables. 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.** Conditioning restricts the sample space using new information.
- **What moves.** Independence is the factorization of joint probability, not the absence of any relationship in ordinary language; Bayes' rule reverses conditioning using base rates.

- **What keeps the answer honest.** The model becomes useful only when I can apply conditional probability, independence, and bayes 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 keep the denominator event visible; most conditional mistakes come from silently changing the reference population.

The easy version

Conditioning restricts the sample space using new information.

Worked example

Question. A disease has 1 percent prevalence, 90 percent sensitivity, and 5 percent false-positive rate. Find positive predictive value.

Walkthrough. Bayes gives 0.009 divided by $0.009 + 0.0495$, approximately 15.4 percent. The low base rate dominates despite good sensitivity.

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. A disease has 1 percent prevalence, 90 percent sensitivity, and 5 percent false-positive rate. Find positive predictive value.

Solution. Bayes gives 0.009 divided by $0.009 + 0.0495$, approximately 15.4 percent. The low base rate dominates despite good sensitivity.

Practice 2. Explain conditional probability, independence, and bayes 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 conditional probability, independence, and bayes 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 conditional probability, independence, and bayes 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 Discrete random variables

The idea

A discrete random variable maps outcomes to countable values. Probability mass functions, cumulative distributions, expectation, variance, Bernoulli, binomial, geometric, negative binomial, hypergeometric, and Poisson models encode different trial structures.

Rebuild the chapter from first principles

The claim I need to own is this: A discrete random variable maps outcomes to countable values. Probability mass functions, cumulative distributions, expectation, variance, Bernoulli, binomial, geometric, negative binomial, hypergeometric, and Poisson models encode different trial structures. 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 discrete random variables. 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 discrete random variables 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 conditional probability, independence, and bayes to continuous random variables. 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 discrete random variable maps outcomes to countable values.
- **What moves.** Probability mass functions, cumulative distributions, expectation, variance, Bernoulli, binomial, geometric, negative binomial, hypergeometric, and Poisson models encode different trial structures.
- **What keeps the answer honest.** The model becomes useful only when I can apply discrete random variables 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 choose a distribution from the data-generating story, not because its formula looks familiar.

The easy version

A discrete random variable maps outcomes to countable values.

Worked example

Question. For X binomial with $n=20$ and $p=0.3$, find mean and variance.

Walkthrough. The mean is $np=6$ and variance is $np(1-p)=4.2$. Independence and constant success probability are part of the model.

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. For X binomial with $n=20$ and $p=0.3$, find mean and variance.

Solution. The mean is $np=6$ and variance is $np(1-p)=4.2$. Independence and constant success probability are part of the model.

Practice 2. Explain discrete random variables 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 discrete random variables 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 discrete random variables 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 Continuous random variables

The idea

Densities integrate to probability and may exceed one at a point; point probability is zero for continuous laws. Uniform, exponential, gamma, beta, and normal distributions model different supports, shapes, and memory structures.

Rebuild the chapter from first principles

The claim I need to own is this: Densities integrate to probability and may exceed one at a point; point probability is zero for continuous laws. Uniform, exponential, gamma, beta, and normal distributions model different supports, shapes, and memory structures. 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 continuous random variables. 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 continuous random variables 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 discrete random variables to joint distributions and random vectors. 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.** Densities integrate to probability and may exceed one at a point; point probability is zero for continuous laws.

- **What moves.** Uniform, exponential, gamma, beta, and normal distributions model different supports, shapes, and memory structures.
- **What keeps the answer honest.** The model becomes useful only when I can apply continuous random variables 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: my shorthand for this chapter is: a density is height, not probability; area is probability.

The easy version

Densities integrate to probability and may exceed one at a point; point probability is zero for continuous laws.

Worked example

Question. If X is exponential with rate 2, find $P(X > 1)$.

Walkthrough. The survival probability is $\exp(-2 \text{ times } 1)$, or about 0.1353. Exponential waiting time is memoryless.

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. If X is exponential with rate 2, find $P(X > 1)$.

Solution. The survival probability is $\exp(-2 \text{ times } 1)$, or about 0.1353. Exponential waiting time is memoryless.

Practice 2. Explain continuous random variables 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 continuous random variables 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 continuous random vari-

ables 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 Joint distributions and random vectors

The idea

Joint, marginal, and conditional laws describe several variables. Covariance and correlation measure linear co-movement but not general dependence; independence implies zero covariance when moments exist, not conversely.

Rebuild the chapter from first principles

The claim I need to own is this: Joint, marginal, and conditional laws describe several variables. Covariance and correlation measure linear co-movement but not general dependence; independence implies zero covariance when moments exist, not conversely. 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 joint distributions and random vectors. 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 joint distributions and random vectors 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 continuous random variables to functions and transformations of variables. 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.** Joint, marginal, and conditional laws describe several variables.
- **What moves.** Covariance and correlation measure linear co-movement but not general dependence; independence implies zero covariance when moments exist, not conversely.
- **What keeps the answer honest.** The model becomes useful only when I can apply joint distributions and random vectors 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: my shorthand for this chapter is: a correlation of zero is a statement about one pattern, not a guarantee the variables ignore each other.

The easy version

Joint, marginal, and conditional laws describe several variables.

Worked example

Question. If $\text{Cov}(X,Y)=3$, $\text{Var}(X)=4$, and $\text{Var}(Y)=9$, find correlation.

Walkthrough. Correlation is 3 divided by $\sqrt{4 \text{ times } 9}$, or 0.5.

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. If $\text{Cov}(X,Y)=3$, $\text{Var}(X)=4$, and $\text{Var}(Y)=9$, find correlation.

Solution. Correlation is 3 divided by $\sqrt{4 \text{ times } 9}$, or 0.5.

Practice 2. Explain joint distributions and random vectors 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 joint distributions and random vectors 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 joint distributions and random vectors 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 Functions and transformations of variables**The idea**

Transformations require mapping support and probability correctly. CDF, change-of-variables, Jacobian, convolution, and order-statistic methods derive distributions of functions, sums, ratios, minima, and maxima.

Rebuild the chapter from first principles

The claim I need to own is this: Transformations require mapping support and probability correctly. CDF, change-of-variables, Jacobian, convolution, and order-statistic methods derive distributions of functions, sums, ratios, minima, and maxima. 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 functions and transformations of variables. 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 functions and transformations of variables 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 joint distributions and random vectors to expectation, conditioning, and generating tools. 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.** Transformations require mapping support and probability correctly.
- **What moves.** CDF, change-of-variables, Jacobian, convolution, and order-statistic methods derive

distributions of functions, sums, ratios, minima, and maxima.

- **What keeps the answer honest.** The model becomes useful only when I can apply functions and transformations of variables 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 draw the transformed support before touching the Jacobian.

The easy version

Transformations require mapping support and probability correctly.

Worked example

Question. If X is uniform on 0 to 1 and $Y=X$ squared, find the CDF of Y on 0 to 1.

Walkthrough. $P(Y \leq y) = P(X \leq \sqrt{y}) = \sqrt{y}$. Differentiating gives density 1 over $2\sqrt{y}$ for $0 < y < 1$.

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. If X is uniform on 0 to 1 and $Y=X$ squared, find the CDF of Y on 0 to 1.

Solution. $P(Y \leq y) = P(X \leq \sqrt{y}) = \sqrt{y}$. Differentiating gives density 1 over $2\sqrt{y}$ for $0 < y < 1$.

Practice 2. Explain functions and transformations of variables 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 functions and transformations of variables 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 functions and transformations of variables 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 Expectation, conditioning, and generating tools

The idea

Linearity of expectation needs no independence. Conditional expectation is a random variable and projection; total expectation and total variance decompose uncertainty. Moment and probability generating functions encode distributions when they exist.

Rebuild the chapter from first principles

The claim I need to own is this: Linearity of expectation needs no independence. Conditional expectation is a random variable and projection; total expectation and total variance decompose uncertainty. Moment and probability generating functions encode distributions when they exist. 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 expectation, conditioning, and generating tools. 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 expectation, conditioning, and generating tools 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 functions and transformations of variables to laws of large numbers and central limit theorem. 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.** Linearity of expectation needs no independence.
- **What moves.** Conditional expectation is a random variable and projection; total expectation and total variance decompose uncertainty.
- **What keeps the answer honest.** Moment and probability generating functions encode distributions when they exist.
- **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

When a direct joint calculation looks ugly, I condition on the variable that makes the rest simple.

The easy version

Linearity of expectation needs no independence.

Worked example

Question. If $E[X|Y]=2Y$ and $E[Y]=3$, find $E[X]$.

Walkthrough. By iterated expectation, $E[X]=E[E[X|Y]]=E[2Y]=6$.

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. If $E[X|Y]=2Y$ and $E[Y]=3$, find $E[X]$.

Solution. By iterated expectation, $E[X]=E[E[X|Y]]=E[2Y]=6$.

Practice 2. Explain expectation, conditioning, and generating tools 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 expectation, conditioning, and generating tools 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 expectation, conditioning, and generating tools 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 Laws of large numbers and central limit theorem

The idea

The law of large numbers makes sample averages concentrate around expectation; the central limit theorem describes scaled fluctuations around that limit. Assumptions, rate, tail behavior, and finite-sample approximation matter.

Rebuild the chapter from first principles

The claim I need to own is this: The law of large numbers makes sample averages concentrate around expectation; the central limit theorem describes scaled fluctuations around that limit. Assumptions, rate, tail behavior, and finite-sample approximation matter. 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 laws of large numbers and central limit theorem. 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 laws of large numbers and central limit theorem 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 expectation, conditioning, and generating tools 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.** The law of large numbers makes sample averages concentrate around expectation; the central limit theorem describes scaled fluctuations around that limit.
- **What moves.** Assumptions, rate, tail behavior, and finite-sample approximation matter.

- **What keeps the answer honest.** The model becomes useful only when I can apply laws of large numbers and central limit theorem 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

LLN tells me where the average goes; CLT tells me the shape of its remaining error.

The easy version

The law of large numbers makes sample averages concentrate around expectation; the central limit theorem describes scaled fluctuations around that limit.

Worked example

Question. A population has mean 10 and standard deviation 4. Approximate the distribution of the mean of 100 iid observations.

Walkthrough. The sample mean is approximately normal with mean 10 and standard error $4/\sqrt{100}=0.4$, assuming the CLT approximation is adequate.

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. A population has mean 10 and standard deviation 4. Approximate the distribution of the mean of 100 iid observations.

Solution. The sample mean is approximately normal with mean 10 and standard error $4/\sqrt{100}=0.4$, assuming the CLT approximation is adequate.

Practice 2. Explain laws of large numbers and central limit theorem 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 laws of large numbers and central limit theorem 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 laws of large numbers and central limit theorem 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.