Discrete Mathematics and Probability

Lecture 19
Conditional and Limit Distributions

Ian Stark

School of Informatics
The University of Edinburgh

Thursday 20 November 2025

Week 11 Topics

Chapter 4: Joint Probability Distributions and Their Applications

- §4.4 Conditional Distributions
- §4.5 The Central Limit Theorem

The study guide and accompanying videos indicate the examinable course content. The corresponding sections in the book are to support this. Additional sections in the book are useful extension material but not required for this course.

Exam Preparation

Next week's lectures on Monday and Thursday will review course content and advise on exam preparation.

Conditional Distribution Functions

Definition

Let X and Y be discrete random variables with joint probability mass function $p(x, y)$ and marginal $p_X(x)$ for X. Then the *conditional probability mass function* (CPMF) of Y given X is defined as follows.

$$p_{Y|X}(y \mid x) = \frac{p(x, y)}{p_X(x)} \qquad \text{for any } x,\ y \text{ where } p_X(x) > 0$$

For continuous random variables X and Y with JPDF $f(x, y)$ and X marginal $f_X(x)$ we have an analogous *conditional probability density function* (CPDF).

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)} \qquad \text{for any } x,\ y \text{ where } f_X(x) > 0$$

Example

A vehicle occupancy survey counts the number of adults and children in passing cars over the course of a week. It finds the following proportions for different combinations of occupancy.

	Children C		
Adults A	0	1	2
1	0.66	0.12	0.05
2	0.14	0.02	0.01

(One value taken from Transport Scotland data; others invented for this example)

For example, the probability that a car chosen at random carries one adult and one child is 0.12.

1. Calculate the conditional probability mass $p_{C|A}(1, 2)$.

Independence

Proposition

gTwo discrete random variables X and Y are independent if and only the conditional PMF of X is the same as its marginal PMF; or similarly for Y.

$$p_{X|Y}(x \mid y) = p_X(x) \iff p(x,y) = p_X(x)p_Y(y) \iff p_{Y|X}(y \mid x) = p_Y(y)$$

The same result holds for continuous random variables and their conditional and marginal probability density functions.

$$f_{X|Y}(x \mid y) = f_X(x) \iff f(x,y) = f_X(x)f_Y(y) \iff f_{Y|X}(y \mid x) = f_Y(y)$$

For independent random variables, conditional probabilities are the same as unconditioned ones.

Proof is direct by algebraic manipulation: see §4.4.1 for details.

Example

A vehicle occupancy survey counts the number of adults and children in passing cars over the course of a week. It finds the following proportions for different combinations of occupancy.

	Children C		
Adults A	0	1	2
1	0.66	0.12	0.05
2	0.14	0.02	0.01

(One value taken from Transport Scotland data; others invented for this example)

For example, the probability that a car chosen at random carries one adult and one child is 0.12.

2 Are the random variables A and C independent?

Conditional Expectation and Variance

Definition

Suppose X and Y are discrete random variables. The *conditional mean* or *conditional expectation* of Y given X is defined from the conditional probability mass $p_{Y|X}(y \mid x)$.

$$\mu_{Y|X=x} \;=\; \mathsf{E}(Y \mid X = x) \;=\; \sum_{y} y \cdot p_{Y|X}(y \mid x)$$

For two continuous random variables X and Y conditional expectation uses integration and the conditional probability density $f_{Y|X}(y \mid x)$.

$$\mu_{Y|X=x} \;=\; \mathsf{E}(Y \mid X = x) \;=\; \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid x)\, dy$$

Example

A vehicle occupancy survey counts the number of adults and children in passing cars over the course of a week. It finds the following proportions for different combinations of occupancy.

	Children C		
Adults A	0	1	2
1	0.66	0.12	0.05
2	0.14	0.02	0.01

(One value taken from Transport Scotland data; others invented for this example)

For example, the probability that a car chosen at random carries one adult and one child is 0.12.

3 Calculate the conditional expectation $\mathsf{E}(C \mid A = 2)$.

Conditional Expectation and Variance

Definition

The *conditional expectation of a function* $h(Y)$ given X for random variables X and Y is defined similarly to the conditional mean.

$$E(h(Y) \mid X = x) \;=\; \begin{cases} \displaystyle\sum_y h(y) \cdot p_{Y|X}(y \mid x) & \text{Discrete random variables } X \text{ and } Y \\ \displaystyle\int_{-\infty}^{\infty} h(y) \cdot f_{Y|X}(y \mid x)\, dy & \text{Continuous random variables } X \text{ and } Y \end{cases}$$

In particular we can calculate the *conditional variance* of Y given X.

$$\sigma^2_{Y|X=x} \;=\; \mathrm{Var}(Y \mid X = x) \;=\; E((Y - \mu_{Y|X=x})^2 \mid X = x) \;=\; E(Y^2 \mid X = x) - \mu^2_{Y|X=x}$$

Example

A vehicle occupancy survey counts the number of adults and children in passing cars over the course of a week. It finds the following proportions for different combinations of occupancy.

	Children C		
Adults A	0	1	2
1	0.66	0.12	0.05
2	0.14	0.02	0.01

(One value taken from Transport Scotland data; others invented for this example)

For example, the probability that a car chosen at random carries one adult and one child is 0.12.

4 Calculate the conditional variance $\text{Var}(C \mid A = 2)$.

Laws of Total Expectation and Variance

Proposition

For random variables X and Y, the conditional mean $E(Y \mid X)$ and variance $Var(Y \mid X)$ of Y given X are themselves both random variables. Each has its own distribution, mean, and variance, with the following properties.

$$E(Y) = E(E(Y \mid X)) \qquad \text{Law of Total Expectation}$$

$$Var(Y) = Var(E(Y \mid X)) + E(Var(Y|X)) \qquad \text{Law of Total Variance}$$

These equations are helpful when the distribution of Y is *only* known by its conditional distribution on X: typically, X describes some environmental factor which can be observed but not controlled and the distribution of Y is given in terms of that.

Example

A vehicle occupancy survey counts the number of adults and children in passing cars over the course of a week. It finds the following proportions for different combinations of occupancy.

	Children C		
Adults A	0	1	2
1	0.66	0.12	0.05
2	0.14	0.02	0.01

(One value taken from Transport Scotland data; others invented for this example)

For example, the probability that a car chosen at random carries one adult and one child is 0.12.

5 Calculate the probability distribution of random variable $\mathrm{E}(C \mid A)$.

Example

A vehicle occupancy survey counts the number of adults and children in passing cars over the course of a week. It finds the following proportions for different combinations of occupancy.

	Children C		
Adults A	0	1	2
1	0.66	0.12	0.05
2	0.14	0.02	0.01

(One value taken from Transport Scotland data;
others invented for this example)

For example, the probability that a car chosen at random carries one adult and one child is 0.12.

6 Calculate the expected value $\mathsf{E}(\mathsf{E}(C \mid A))$.

Week 11 Topics

Chapter 4: Joint Probability Distributions and Their Applications

§4.4 Conditional Distributions

§4.5 The Central Limit Theorem

The study guide and accompanying videos indicate the examinable course content. The corresponding sections in the book are to support this. Additional sections in the book are useful extension material but not required for this course.

Exam Preparation

Next week's lectures on Monday and Thursday will review course content and advise on exam preparation.

Random Samples

Definition

A set of random variables $X_1, X_2, \ldots, X_n$ are *independent and identically distributed* (IID) if:

- The random variables X_i are all independent; and
- Every X_i has the same probability distribution.

We call such a set of random variables a *random sample* of size n from this distribution.

Total and Mean

For a random sample $X_1, \ldots, X_n$ of size n the *sample total* T and *sample mean* $\bar{X}$ are two random variables defined from the X_i.

$$T = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i \qquad \bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{T}{n}$$

Properties of Sample Total and Mean

Proposition

Let T and $\bar{X}$ be sample total and mean of IID random variables $X_1, \ldots, X_n$ with mean μ and standard deviation σ. Then they have the following properties.

- $E(T) = n\mu$
- $\mathrm{Var}(T) = n\sigma^2$ and $SD(T) = \sqrt{n}\sigma$
- If the X_i are normally distributed, then so is T.

- $E(\bar{X}) = \mu$
- $\mathrm{Var}(\bar{X}) = \sigma^2/n$ and $SD(\bar{X}) = \sigma/\sqrt{n}$
- If the X_i are normally distributed, then so is $\bar{X}$.

Proofs use previous results on linear combinations of random variables: that expectations always add up; variances do too if the random variables are independent; and independent normal distributions add to give normal distributions.

The Law of Large Numbers

Proposition

Let $X_1, \ldots, X_n$ be a random sample from a distribution where each X_i has mean μ and standard deviation σ. In the limit as $n \to \infty$ the sample mean $\bar{X}$ converges to μ:

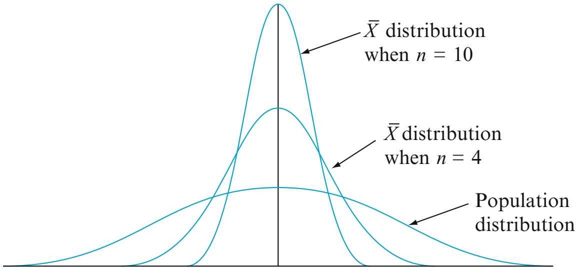
- In mean square: $E\big((\bar{X} - \mu)^2\big) \to 0$ as $n \to \infty$
- In probability: $P(|\bar{X} - \mu| \geqslant \varepsilon) \to 0$ as $n \to \infty$ for any $\varepsilon > 0$.

This applies for any random sample whatever the original distribution: the sample mean always converges to the distribution mean. Finally, an "expected value" really is something expected.

Sampling Normal Distributions

Let $X_1, \ldots, X_n$ be a random sample of normally-distributed variables, each with mean μ and standard deviation σ. Then $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$.

$$\begin{array}{rl} \text{If} & X_i \sim N(\mu, \sigma) \\ \text{Then} & \bar{X} \sim N(\mu, \sigma/\sqrt{n}) \end{array}$$

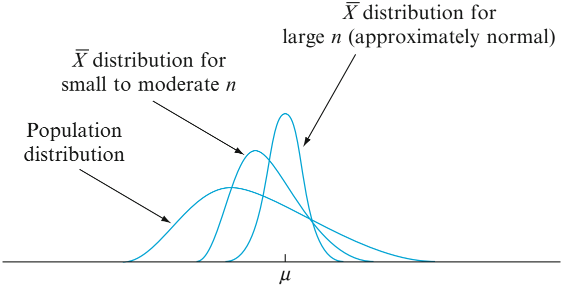


Carlton & Devore Figure 4.8

Sampling Arbitrary Distributions

Let $X_1, \dots, X_n$ be a random sample of IID variables, each with mean μ and standard deviation σ. Then as n becomes large the random variable $\bar{X}$ approaches a normal distribution.

$$\text{“}\lim_{n\to\infty} \bar{X} \sim N(\mu, \sigma/\sqrt{n})\text{”}$$



Carlton & Devore Figure 4.10

The Central Limit Theorem

Proposition

Let $X_1, \dots, X_n$ be a random sample from a distribution where each X_i has mean μ and standard deviation σ. In the limit as $n \to \infty$ the sample total T and sample mean $\bar{X}$ have normal distributions.

$$\lim_{n\to\infty} P\left(\frac{T - n\mu}{\sqrt{n}\sigma} \leqslant z\right) = P(Z \leqslant z) = \Phi(z) \qquad \text{“}\lim_{n\to\infty} T \sim N(n\mu, \sqrt{n}\sigma)\text{”}$$

$$\lim_{n\to\infty} P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leqslant z\right) = P(Z \leqslant z) = \Phi(z) \qquad \text{“}\lim_{n\to\infty} \bar{X} \sim N(\mu, \sigma/\sqrt{n})\text{”}$$

Here $Z \sim N(0,1)$ is a standard normal variable. We say that random variables T and $\bar{X}$ are *asymptotically normal*.

Carcinisation


Wikipedia: Bhny


Thom Quine


Museums Victoria


Martin-Lefèvre P.


Epic Wildlife


J. Antonio Baeza

Summary

Topics

- Conditional distribution functions: discrete and continuous
- Independence
- Conditional expectation and variance
- Laws of total expectation and variance
- Random samples
- Sample total and mean
- Sampling normal distributions
- Central Limit Theorem
- Law of Large Numbers

Reading

Chapter 4, §4.4, §4.5, §§4.5.1, 4.5.2, and 4.5.4; pp. 277–286, 290–297 and 299.

Exercises

Chapter 4, Exercises 66–100, pp. 286–289 and 300–302.