Session 09-04 - Probability & Statistics Review

Section 09: Exam Preparation

Dr. Nikolai Heinrichs & Dr. Tobias Vlćek

Warm-Up Quiz - 10 Minutes

Quick Review Questions

Work individually for 5 minutes, then discuss with the class (5 minutes)

1. In how many ways can you choose 3 items from 10?
2. Events A and B are independent with $P(A) = 0.4$ and $P(B) = 0.3$. Find $P(A \cup B)$.
3. If $X \sim B(10, 0.2)$, find $E[X]$ and $\text{Var}\ (X)$.
4. An investment of €$10,000$ earns 4% compound interest. What is it worth after 5 years?

Guided Practice - Part A - 25 Minutes

Practice Problem 1

Committee Selection (Section 07)

Work in pairs

A committee of 5 is to be formed from 8 men and 6 women.

(a) How many committees are possible?

(b) How many committees have exactly 3 women?

(c) What is the probability that exactly 3 women are selected?

(d) What is the probability of at least 1 woman?

Practice Problem 2

Bayes' Theorem & Contingency Table (Section 07)

Work in pairs

A medical test has sensitivity (true positive rate) of 95% and specificity (true negative rate) of 90%. In a population with disease prevalence of 2%.

(a) Compute the Positive Predictive Value using Bayes' theorem.

(b) Construct a contingency table for 10,000 people.

(c) Verify your answer from (a) using the table.

(d) Compute the Negative Predictive Value.

Practice Problem 3

System of Equations from Probability (Sections 02 & 07)

Work in pairs

For events A and B: $P(A) + P(B) = 0.7$, $P(A \cap B) = 0.1$, and $P(A \mid B) = 0.25$.

(a) Use $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ to find $P(B)$.

(b) Solve the system for $P(A)$.

(c) Are A and B independent?

(d) Find $P(A \cup B)$.

Practice Problem 4

Binomial Distribution (Section 07)

Work in pairs

A quality inspector tests 20 items from a production line where the defect rate is 5%.

(a) What distribution models this? State parameters.

(b) Find $P(X = 0)$ (no defects).

(c) Find $P(X \leq 2)$.

(d) Find $E[X]$ and σ.

(e) How many items should be tested so that $P(X \geq 1) > 0.99$?

Practice Problem 5

Descriptive Statistics (Section 07)

Work in pairs

Employee monthly salaries (in €) at a small company:

Salary Range	2000–3000	3000–4000	4000–5000	5000–6000	6000–8000
Frequency	8	15	12	5	2

(a) Estimate the mean salary.

(b) Find the median class.

(c) Calculate the standard deviation.

(d) The company claims “average salary is over €$4,000$.” True?

Guided Practice - Part B - 30 Minutes

Practice Problem 6

Hypergeometric vs Binomial (Section 07)

Work in pairs

A box has 15 items, 3 of which are defective. You draw 5 items without replacement.

(a) Find $P(X = 1)$ using the hypergeometric distribution.

(b) Compare with the binomial approximation $B(5, 0.2)$.

(c) When is the binomial a good approximation for the hypergeometric?

Practice Problem 7

Compound Interest + Probability (Sections 08 & 07)

Work in pairs

You invest $€1,000$ per year for 10 years. Each year, the interest rate is 5% with probability 0.7 and 2% with probability 0.3 (independent across years).

(a) What is the expected interest rate?

(b) What is the expected value after 10 years using the expected rate?

(c) Compute the actual expected value after 1 year if the interest rate is 5% with probability 0.7 and 2% with probability 0.3. Compare with the result using the expected rate from (b). Are they the same?

Practice Problem 8

Two-Way Independence Test (Section 07)

Work in pairs

In a study of 400 employees: 200 work full-time and 200 part-time. Among full-time workers: 140 are satisfied and 60 are dissatisfied. Among part-time workers: 120 are satisfied and 80 are dissatisfied.

(a) Construct the contingency table.

(b) Are satisfaction and work type independent? Compute $P(\text{Satisfied} \mid \text{Full-time})$ vs $P(\text{Satisfied})$.

(c) Compute the expected frequencies under independence.

(d) Based on your findings, are work type and satisfaction independent? Justify using conditional probabilities.

Coffee Break - 15 Minutes

Collaborative Problem-Solving - Part A - 25 Minutes

Practice Problem 9

Insurance & Probability (Sections 07 & 08)

An insurance company covers 1,000 policyholders. Each pays $€200$/month ($€2,400$/year). Each policyholder has a 3% probability of filing a claim worth $€15,000$ per year (independent events).

(a) What is the expected number of claims?

(b) What is the expected total payout?

(c) What is the expected profit?

(d) What is the probability that the company pays out more than it collects? (i.e., $P(X > 160)$ where X is the number of claims.)

(e) What reserve should the company hold to be 95% confident of covering all claims?

Practice Problem 10

Logarithms & Binomial (Sections 07 & 01)

Think individually (2 min), then work in groups of 3-4

A machine has a probability of $p = 0.02$ of failing on any given day (independent).

(a) What is the probability that the machine runs for at least 30 days without failure?

(b) Solve for n: how many days until the probability of at least one failure exceeds 99%? (Solve $(1 - 0.02)^n < 0.01$ using logarithms.)

(c) If a backup machine costs $€500$/day and downtime costs $€10,000$, is it cost-effective to keep a backup permanently?

Practice Problem 11

Multi-Step Bayes' Theorem (Section 07)

A factory has three production lines:

- Line 1 produces 50% of output with 2% defect rate
- Line 2 produces 30% of output with 3% defect rate
- Line 3 produces 20% of output with 5% defect rate

(a) A randomly selected item is defective. Probability it came from Line 3?

(b) Construct a contingency table for 1,000 items.

(c) Two items are randomly selected. Probability both are defective?

(d) Given at least one of two is defective. Probability both are defective?

Coffee Break - 10 Minutes

Collaborative Problem-Solving - Part B - 25 Minutes

Practice Problem 12

Reliability Engineering (Sections 07 & 01)

Think individually (2 min), then work in groups of 3-4

A system has 3 components. Each fails independently with probabilities 0.05, 0.03, and 0.02.

(a) If the components are in series (system works only if ALL work), what is the probability the system works?

(b) If the components are in parallel (system works if ANY works), what is the reliability?

(c) How many identical components with $p = 0.05$ failure rate must be placed in parallel to achieve 99.9% reliability? (Use logarithms!)

Practice Problem 13

Bayesian Updating with Data (Section 07)

Think individually (2 min), then work in groups of 3-4

A coin is either fair ($p = 0.5$) or biased ($p = 0.8$). Your prior belief: 50% chance each. You flip the coin 5 times and observe: H, H, H, T, H (4 heads, 1 tail).

(a) Compute $P(\text{data} \mid \text{fair})$ and $P(\text{data} \mid \text{biased})$.

(b) Update the probability the coin is biased.

(c) If you flip once more and get H, update again.

Practice Problem 14

Expected Value Decision Problem (Sections 07 & 08)

A company must decide between two investments:

- Project: Invest $€50,000$. With probability 0.6, it succeeds and returns $€120,000$. With probability 0.4, it fails and returns $€10,000$.
- Bonds: Invest $€50,000$ at a guaranteed 4% annual interest for 2 years.

(a) Compute the expected return of the project.

(b) Compute the bond return after 2 years.

(c) Compute the variance of the project return.

(d) Which should a risk-averse company choose? Discuss.

Wrap-Up & Next Steps

Key Takeaways

- Combinatorics & probability — apply Bayes' theorem and verify independence via conditional probabilities
- Binomial & hypergeometric — use binomial approximation when sample $\leq$ 5% of population
- Descriptive statistics — use midpoints for grouped data; compare mean vs median for skew
- Log equations in probability — flip inequalities when dividing by negative log values
- Decisions under uncertainty — consider variance and risk aversion, not just expected value