Tasks 07-05 - Bayes' Theorem

Section 07: Probability & Statistics

Problem 1: Basic Bayes' Formula (x)

In a factory, 60% of products come from Machine A and 40% from Machine B. Machine A has a 2% defect rate, while Machine B has a 5% defect rate.

a) What is the probability that a randomly selected product is defective?
b) If a product is defective, what is the probability it came from Machine A?

Problem 2: Medical Testing Basics (x)

A disease affects 1% of the population. A screening test has: - Sensitivity (true positive rate): 95% - Specificity (true negative rate): 90%

a) If a person tests positive, what is the probability they have the disease (PPV)?
b) If a person tests negative, what is the probability they don't have the disease (NPV)?

Problem 3: Quality Control (xx)

A company has three suppliers: - Supplier A: 50% of parts, 4% defect rate - Supplier B: 30% of parts, 2% defect rate - Supplier C: 20% of parts, 6% defect rate

a) What is the overall defect rate?
b) A part is found defective. What is the probability it came from Supplier C?
c) Which supplier should be investigated first for quality issues?

Problem 4: Two-Stage Testing (xx)

A rare disease affects 0.5% of the population. A two-stage testing protocol: - Stage 1 test: 98% sensitivity, 85% specificity - Stage 2 test (only if Stage 1 positive): 99% sensitivity, 95% specificity

a) What is the probability of testing positive in Stage 1?
b) Given a positive Stage 1 result, what is the probability of having the disease?
c) Given positive results in both stages, what is the probability of having the disease?

Problem 5: Email Spam Filter (xx)

A spam filter classifies emails. From historical data: - 30% of emails are spam - When an email is spam, the filter correctly identifies it 92% of the time - When an email is not spam, the filter incorrectly marks it as spam 8% of the time

a) What percentage of all emails are marked as spam by the filter?
b) If an email is marked as spam, what is the probability it's actually spam?

c) If an email passes the filter (not marked as spam), what is the probability it's actually not spam?

Problem 6: Effect of Prevalence on PPV (xxx)

A test has sensitivity 95% and specificity 90%. Calculate the PPV for three different prevalence rates:

a) Prevalence = 10% (common condition)
b) Prevalence = 1% (uncommon condition)
c) Prevalence = 0.1% (rare condition)
d) What pattern do you observe? Explain why this matters for screening.

Problem 7: What Prevalence is Needed? (xxx)

A new diagnostic test has 90% sensitivity and 95% specificity.

a) What disease prevalence would be needed for PPV = 50%?
b) What disease prevalence would be needed for PPV = 80%?
c) What disease prevalence would be needed for PPV = 95%?

Problem 8: Insurance Risk Assessment (xxx)

An insurance company categorizes drivers: - 20% are "high risk" (accident probability 15%) - 50% are "medium risk" (accident probability 5%) - 30% are "low risk" (accident probability 1%)

a) What is the overall accident probability for a randomly selected driver?
b) A new customer has an accident in their first year. What is the probability they are high risk?
c) The company wants to assign risk categories based on accident history. After 3 accident-free years, what is the probability a driver is low risk? (Assume independence)

Problem 9: Serial Testing (xxx)

A disease has 5% prevalence. Two independent tests are available: - Test A: 90% sensitivity, 85% specificity - Test B: 95% sensitivity, 80% specificity

Compare two testing strategies:

a) Strategy 1: Use only Test A. What is the PPV?
b) Strategy 2: First use Test B, then if positive, confirm with Test A. What is the final PPV?
c) Which strategy is better? Why might a healthcare system choose the worse PPV strategy?

Problem 10: Exam-Style Problem (xxxx)

A company produces electronic components. Components are manufactured by three machines: - Machine 1: produces 40% of output, 3% defect rate - Machine 2: produces 35% of output, 2% defect rate - Machine 3: produces 25% of output, 5% defect rate

An automatic inspection system tests each component: - If a component is defective, the system detects it with 95% probability - If a component is good, the system incorrectly flags it as defective with 8% probability

a) What is the overall probability that a randomly selected component is defective?
b) What is the probability that a component flagged by the inspection system is actually defective?
c) Given that a component passed inspection, what is the probability it came from Machine 2?
d) The company wants to improve quality. If they eliminate Machine 3, what would be the new overall defect rate?

Problem 11: Tree Diagram Analysis (xxx)

Draw a probability tree and solve:

A marketing campaign targets customers. Historical data shows: - 40% of customers receive email marketing - 30% of customers receive phone marketing - 30% of customers receive both

For customers who receive email marketing only: 10% purchase For customers who receive phone marketing only: 15% purchase For customers who receive both: 25% purchase For customers who receive neither: 2% purchase

a) What is the overall purchase rate?
b) If a customer made a purchase, what is the probability they received both types of marketing?
c) Which marketing channel is most effective in terms of increasing purchase probability?

Problem 12: Comprehensive Medical Testing (xxxx)

A hospital screens for a disease with the following characteristics: - Disease prevalence in the screening population: 2% - Screening test: 92% sensitivity, 88% specificity - Confirmatory test: 99% sensitivity, 97% specificity

Testing protocol: All patients get the screening test. Those who test positive get the confirmatory test.

a) What percentage of the screening population will need the confirmatory test?
b) Of those who test positive on the screening test, what percentage will test positive on the confirmatory test?
c) If a patient tests positive on both tests, what is the probability they have the disease?
d) The hospital can only afford confirmatory tests for 10% of patients. What specificity would the screening test need to achieve this?