STAT 131: Practice Exam

Introductory Probability Theory

Prepared by Antonio Aguirre

Problem 1 – Discrete Random Variables & Expectation

Let X be the number of defective items in a random sample of 3 items taken *without replacement* from a batch of 10 items, of which 2 are defective and 8 are good.

(a) Write down the **support** of X and the corresponding probability mass function $P(X = x)$.

(b) Compute $\mathbb{E}[X]$ directly from the pmf.

(c) Argue why $\mathbb{E}[X]$ must equal the expected number of defectives in 3 draws using a **linearity-of-expectation / indicator** argument (without referencing the pmf).

(d) Compute $\mathrm{Var}(X)$.

(e) **(Topic: approximation)** Suppose now the batch is much larger: 2 defective in 1000 items, and you still sample 3 without replacement.

  (i) Explain why a **Binomial** approximation is reasonable here.

  (ii) Write the approximate distribution of X and recompute $\mathbb{E}[X]$ and $\mathrm{Var}(X)$ under this approximation.

Problem 2 – Joint & Conditional Distributions (Discrete)

Two fair six-sided dice are rolled. Let

$$X = \text{the sum of the two dice}, \qquad Y = \text{the maximum of the two dice}.$$

(a) List all possible pairs (X, Y) and identify the **support** of the joint distribution (X, Y).

(b) Compute $P(X = 7, Y = 5)$.

(c) Find the **marginal pmf** of Y, i.e. $P(Y = y)$ for all possible y.

(d) Compute the conditional probability $P(X = 7 \mid Y = 5)$.

(e) **(Topic: independence)** Are X and Y independent? Justify using the definition of independence in terms of the joint and marginal distributions.

(f) **(Topic: covariance)** Compute $\mathrm{Cov}(X, Y)$. Interpret the sign of the covariance in words.

Problem 3 – Continuous Joint Density & Conditioning

Let (X, Y) have joint density

$$f_{X,Y}(x, y) = c(x + y), \qquad 0 \le x \le 1,\ 0 \le y \le 1,$$

and $f_{X,Y}(x, y) = 0$ otherwise.

- **(a)** Find the normalizing constant c.
- **(b)** Compute the **marginal densities** $f_X(x)$ and $f_Y(y)$.
- **(c)** Compute the conditional density $f_{Y|X}(y \mid x)$.
- **(d)** For a fixed x, identify the **distribution type** of $Y \mid X = x$. Briefly justify.
- **(e)** **(Topic: expectation via conditioning)** Use the conditional density to compute $\mathbb{E}[Y \mid X = x]$, and then use the **law of total expectation** to find $\mathbb{E}[Y]$.
- **(f)** **(Topic: independence check)** Are X and Y independent? Give a formal argument using the relationship between $f_{X,Y}$, f_X, and f_Y.

Problem 4 – Law of Total Probability & Bayes (Applied)

A medical test is designed to detect a certain disease. In the population:

- The disease prevalence is $P(D) = 0.01$.
- The test has **sensitivity** $P(\text{Pos} \mid D) = 0.95$.
- The test has **specificity** $P(\text{Neg} \mid D^c) = 0.98$.

Here D denotes disease status, and Pos and Neg denote positive and negative test results.

(a) Compute $P(\text{Pos})$.

(b) Compute $P(D \mid \text{Pos})$.

(c) Interpret the value from part (b) in plain language.

(d) **(Topic: effect of prevalence)** Suppose in a high-risk subgroup the prevalence is $P(D) = 0.10$. Recompute $P(D \mid \text{Pos})$.

Problem 5 – Central Limit Theorem & Approximation

A factory produces screws whose lengths are i.i.d. with mean $\mu = 5$ cm and standard deviation $\sigma = 0.2$ cm. Let $\bar{X}_n$ be the sample mean length of n screws.

(a) For $n = 50$, approximate

$$P\big(|\bar{X}_{50} - 5| \leq 0.05\big)$$

using the Central Limit Theorem.

(b) For $n = 200$, approximate the same probability.

(c) Compare your answers and explain qualitatively why the probability changes as n increases.

(d) **(Topic: solving for n)** Solve for the smallest integer n such that

$$P\big(|\bar{X}_n - 5| \leq 0.05\big) \approx 0.99.$$

(e) **(Conceptual)** Explain the CLT in this context in 3–4 sentences.

Problem 6 – Conditioning, Total Expectation, Total Variance & MGF Practice

A random variable Z takes values in $\{0, 1\}$ with

$$P(Z = 1) = 0.3, \qquad P(Z = 0) = 0.7.$$

Conditional on Z, a continuous random variable X is defined as:

$$X \mid Z = 0 \sim \text{Uniform}(0, 1), \qquad X \mid Z = 1 \sim \text{Uniform}(1, 3).$$

(a) Compute $\mathbb{E}[X \mid Z]$ and $\text{Var}(X \mid Z)$ for each value of Z.

(b) Use the **Law of Total Expectation** to compute:

$$\mathbb{E}[X] = \mathbb{E}\big(\mathbb{E}[X \mid Z]\big).$$

(c) Use the **Law of Total Variance**:

$$\text{Var}(X) = \mathbb{E}[\text{Var}(X \mid Z)] + \text{Var}(\mathbb{E}[X \mid Z])\,.$$

(d) **MGF practice.** For a Uniform(a, b) distribution, recall its moment generating function:

$$M_X(t) = \frac{e^{tb} - e^{ta}}{t(b - a)}.$$

(i) Write the MGF of $X \mid Z = 0$ and of $X \mid Z = 1$.

(ii) Compute $M_X'(0)$ for each conditional distribution to recover

$$\mathbb{E}[X] = M_X'(0).$$

(iii) Briefly explain why this matches your answer from part (b).

(e) **(Interpretation)** Explain in 2–3 sentences how conditioning simplifies the computation of expectations and variances in mixture-type models.