

# MT1 Review<sup>1</sup>

Math-UA 185 Prob Stat

Ed Chen

Feb 20, 2026

---

<sup>1</sup>Slides at <https://cims.nyu.edu/~ehc7681/185>

- ▶ Combinatorics, Counting Problems
- ▶ Conditional Probability
- ▶ Independence
- ▶ Discrete RVs
- ▶ Continuous RVs
- ▶ Expectation, Variance

Topics are exhaustive, but this slideset is not!

- ▶ Review past exams, homework, quizzes.

## Counting Rules

- ▶ (Product Rule) If two disjoint tasks have  $m$  **and**  $n$  choices, there are  $mn$  choices.
- ▶ (Sum Rule) If a single task has to set of outcomes  $m$  **or**  $n$  choices, there are  $m + n$  choices.

Q: How many integers less than 500 end in 0?

## Counting Rules

- ▶ (Product Rule) If two disjoint tasks have  $m$  **and**  $n$  choices, there are  $mn$  choices.
- ▶ (Sum Rule) If a single task has to set of outcomes  $m$  **or**  $n$  choices, there are  $m + n$  choices.

Q: How many integers less than 500 end in 0?

A: **Sum** 3 sets of mutually exclusive outcomes  $\{1, 2, 3\}$  digits

**Product** for  $n$  digit numbers, there are  $n$  disjoint "subtasks"

$$1 + (9)(1) + (4)(10)(1) = 50$$

## Sampling

Say we want to choose  $k$  items from  $n$ . If the sampling is...  
*without replacement*

- ▶ and order matters:  $\frac{n!}{(n-k)!}$
- ▶ and order doesn't matter  $\binom{n}{k}$

*with replacement*

- ▶ and order matters:  $n^k$
- ▶ and order doesn't matter<sup>a</sup>:  $\binom{n+k-1}{n}$  types

---

<sup>a</sup>The intuitive reasoning why this is stars and bars, is the unordered outcome is equivalent to saying you have  $n$  of each thing, equivalently you can portion the  $k$  items into  $n$

Q: How many distinct permutations are there of the word "SUCCESSFULLY" ?

Q: How many 5 card hands of a standard 52 card deck have exactly two hearts?

Q: How many distinct permutations are there of the word "SUCCESSFULLY" ?

A: 3 Ss, 2 Us, 2 Cs, 2 Ls that are not distinct. So

$$\frac{12!}{3!2!2!2!}$$

Q: How many 5 card hands of a standard 52 card deck have exactly two hearts? Choose 2 of 13 hearts, and then choose 3 non-hearts from 52-13 non-hearts

$$\binom{13}{2} \binom{39}{3}$$

**Important**

- ▶ **Range:** For any event  $A$ ,  $0 \leq P(A) \leq 1$
- ▶ **Complement:**  $P(A^C) = 1 - P(A)$
- ▶ **Monotonicity:** If  $A \subseteq B$  then  $P(A) \leq P(B)$ 
  - ▶ Consequentially for any  $B$ ,  $P(A \cap B) \leq P(A)$
- ▶ **DeMorgan's Laws**
  - ▶  $(A \cup B)^C = A^C \cap B^C$  and  $(A \cap B)^C = A^C \cup B^C$
- ▶ **Normalization:** For the universal set  $\Omega$ ,  $P(\Omega) = 1$
- ▶ **Total Prob:**  $P(A) = P(A \cap B) + P(A \cap B^C)$

## Inclusion/Exclusion

Adding probabilities of non mutually exclusive events, overlaps get overcounted and need to be accounted for. For two events  $A, B$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

For events  $A_1, \dots, A_n$ ,

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_i P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots$$

This also holds for counting.

Q: How many integers in  $\{1, 2, \dots, 1000\}$  are divisible by 4 **or** 6?

Q: How many integers in  $\{1, 2, \dots, 1000\}$  are divisible by 4 **or** 6?

A: Let  $A_4 = \{n : 4 \mid n\}$  and  $A_6 = \{n : 6 \mid n\}$ . Then

$$|A_4 \cup A_6| = |A_4| + |A_6| - |A_4 \cap A_6|.$$

$$|A_4| = \left\lfloor \frac{1000}{4} \right\rfloor = 250, \quad |A_6| = \left\lfloor \frac{1000}{6} \right\rfloor = 166,$$

$$|A_4 \cap A_6| = |\{n : \text{lcm}(4, 6) \mid n\}| = \left\lfloor \frac{1000}{12} \right\rfloor = 83.$$

$$|A_4 \cup A_6| = 250 + 166 - 83 = 333.$$

## Conditional Probability

If  $P(B) > 0$  the conditional probability  $P(A|B)$  describes the probability that  $A$  happens **given**  $B$  happened. Equivalently,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

## Total Probability

If disjoint events  $B_i$  span the universal set  $\bigcup_i B_i = \Omega$ , then

$$P(A) = \sum_i P(A|B_i)P(B_i)$$

Commonly, in  $n = 2$ ,  $P(A) = P(A|B)P(B) + P(A|B^C)P(B^C)$

## Independence

Two events are **independent** if

$$P(A \cap B) = P(A)P(B)$$

A collection of events is **pairwise independent** if for every distinct  $i, j \in S$

$$P(A_i \cap A_j) = P(A_i)P(A_j)$$

**mutually independent** if for every subset  $S$

$$P\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} P(A_i)$$

Q: A, B are independent. B, C are independent. A, C are independent. Can we say A,B,C are mutually independent?

Q: Can two mutually exclusive events with nonzero probability be independent?

Q: A, B are independent. B, C are independent. A, C are independent. Can we say A,B,C are mutually independent?

A: No. The subset of 3 events  $P(A, B, C) = P(A)P(B)P(C)$  does not follow.

Q: Can two mutually exclusive events with nonzero probability be independent?

A: No. We are given  $P(A \cap B) = 0$ ,  $P(A) > 0$ , and  $P(B) > 0$ . Necessarily  $P(A)P(B) > 0$  which contradicts the first given.

Q:  $P(B) = 0.3$ ,  $P(A \cap B) = 0.1$ ,  $P(A|B^C) = 0.1$ . What is  $P(A)$ ?

Q:  $P(B) = 0.3$ ,  $P(A \cap B) = 0.1$ ,  $P(A|B^C) = 0.1$ . What is  $P(A)$ ?

A: Expand total probability, then conditional probability

$$\begin{aligned}P(A) &= P(A|B)P(B) + P(A|B^C)P(B^C) \\&= \frac{P(A \cap B)}{P(B)}P(B) + P(A|B^C)(1 - P(B)) \\&= 0.1 + 0.1(1 - 0.3) = 0.17\end{aligned}$$

Are A and B independent? No,  $P(A \cap B) \neq P(A)P(B)$

Q:  $P(B) = 0.4$ ,  $P(A \cap B) = 0.1$ ,  $P(A|B^C) = 0.1$ . Find  $P(B|A^C)$ ?

Q:  $P(B) = 0.4$ ,  $P(A \cap B) = 0.1$ ,  $P(A|B^C) = 0.1$ . Find  $P(B|A^C)$ ?

A: Note that by total probability

$$P(B \cap A^C) = P(B) - P(B \cap A) = 0.3$$

$$P(A) = P(A \cap B) + P(A \cap B^C) = P(A \cap B) + P(A|B^C)P(B^C) = 0.16$$

Q:  $P(B) = 0.4$ ,  $P(A \cap B) = 0.1$ ,  $P(A|B^C) = 0.1$ . Find  $P(B|A^C)$ ?

A: Note that by total probability

$$P(B \cap A^C) = P(B) - P(B \cap A) = 0.3$$

$$P(A) = P(A \cap B) + P(A \cap B^C) = P(A \cap B) + P(A|B^C)P(B^C) = 0.16$$

Then, by conditional probability definition

$$\begin{aligned} P(B|A^C) &= \frac{P(B \cap A^C)}{P(A^C)} = \frac{P(B \cap A^C)}{1 - P(A)} \\ &= \frac{0.3}{0.84} \end{aligned}$$

**Bayes**

For events  $A, B$  with  $P(A) > 0$  and  $P(B) > 0$ ,

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

for disjoint events  $A_i$  span the universal set  $\bigcup_i A_i = \Omega$

$$P(A_j | B) = \frac{P(B | A_j)P(A_j)}{\sum_i P(B | A_i)P(A_i)}.$$

Generally  $P(A|B) \neq P(B|A)$ !

Q: A disease occurs in 2% of a population. A test has sensitivity 0.95 (it is positive with probability 0.95 if the person has the disease) and specificity 0.90 (it is negative with probability 0.90 if the person does not have the disease). A person tests positive. What is  $P(\text{disease} \mid +)$ ?

Q: A disease occurs in 2% of a population. A test has sensitivity 0.95 (it is positive with probability 0.95 if the person has the disease) and specificity 0.90 (it is negative with probability 0.90 if the person does not have the disease). A person tests positive. What is  $P(\text{disease} | +)$ ?

A:

$$\begin{aligned} P(D|+) &= \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)} \\ &= \frac{0.95 \times 0.02}{0.95 \times 0.02 + (1 - 0.90)(1 - 0.02)} \end{aligned}$$

### Discrete RV

A random variable  $X$  is **discrete** if it takes values in a countable set (e.g.  $\{0, 1, 2, \dots\}$ ), and probabilities are given by a PMF.

### PMF/CDF

A PMF describes the probability any value happens

$$p_X(x) = P(X = x)$$

and satisfies the properties  $p_X(x) \geq 0, \sum_x p_X(x) = 1$ . A CDF is defined as

$$F_X(x) = P(X \leq x) = \sum_{t \leq x} p_X(t).$$

### Bernoulli( $p$ )

$X \sim \text{Bern}(p)$  means  $P(X = 1) = p$  and  $P(X = 0) = 1 - p$ .

### Binomial( $n, p$ )

$X \sim \text{Bin}(n, p)$  counts the number of successes in  $n$  independent trials:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n.$$

Binomial  $\approx$  “add up  $n$  Bernoullis.”

### Geometric( $p$ ) (1st success time)

$T \sim \text{Geom}(p)$  means

$$P(T = k) = (1 - p)^{k-1}p, \quad k = 1, 2, 3, \dots$$

( $T$  = number of trials until the first success.)

### Poisson( $\lambda$ )

$X \sim \text{Pois}(\lambda)$ :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

(Counts “rare” events in a fixed interval.)

Q: A basketball player makes each free throw with probability  $p$ , independently. Let  $T$  be the number of free throws until the first made shot. Find  $P(T > 5)$ .

Q: Emails arrive according to a Poisson process at rate  $\lambda = 3$  per hour. What is the probability you get at least 2 emails in 30 minutes?

Q: A basketball player makes each free throw with probability  $p$ , independently. Let  $T$  be the number of free throws until the first made shot. Find  $P(T > 5)$ .

A:  $T \sim \text{Geom}(p)$ . " $T > 5$ " means first 5 are misses:

$$P(T > 5) = (1 - p)^5.$$

Q: Emails arrive according to a Poisson process at rate  $\lambda = 3$  per hour. What is the probability you get at least 2 emails in 30 minutes?

A: In 0.5 hours, mean is  $\lambda' = 3(0.5) = 1.5$ , so  $X \sim \text{Pois}(1.5)$ .

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-1.5} - e^{-1.5}(1.5).$$

## Continuous RV

A random variable  $X$  is **continuous** if there is a function  $f_X$  such that for any **closed interval**  $[a, b]$ ,

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx.$$

**Normalization:**  $f_X(x) \geq 0$  for all  $x$ , and

$$\int_{-\infty}^{\infty} f_X(x) dx = 1.$$

**Massless:** For a continuous RV,

$$P(X = c) = 0 \quad \text{for any constant } c.$$

## CDFs for continuous RVs

For continuous RVs, the CDF definition is

$$F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

recall you can find the probability on the **open interval**  $(a,b]$  with

$$P(a < X \leq b) = F_X(b) - F_X(a).$$

If it is differentiable, the relation to PDF is given by the FTC.

$$f_X(x) = F'_X(x),$$

**Definition**

$X \sim \text{Unif}(a, b)$  means  $X$  is equally likely anywhere in  $[a, b]$ .

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

Example: Bus arrives every 30 minutes, waiting time when you arrive randomly.

**Definition**

$X \sim \text{Exp}(\lambda)$  models a **waiting time** with rate  $\lambda > 0$ .

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

**Memoryless** distribution, that  $P(X > s + t | X > s) = P(X > t)$   
Example: Lifespan of a lightbulb. The implicit assumption is that the lifespan of a lightbulb is not related to the age of the lightbulb, it breaking is a random event that has a constant chance at any given time.

**Definition**

$X \sim N(\mu, \sigma^2)$  is a **bell-shaped** continuous distribution with mean  $\mu$  and variance  $\sigma^2$ .

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$$

If  $X \sim N(\mu, \sigma^2)$  then

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Example: Heights of people

Q: Let  $X \sim \text{Exp}(0.2)$ . Find  $P(X > 10)$ .

Q: Find  $P(3 < X \leq 8)$ .

Q: Let  $X \sim \text{Exp}(0.2)$ . Find  $P(X > 10)$ .

A: Use the tail from the CDF:

$$P(X > 10) = 1 - F_X(10) = e^{-0.2 \cdot 10} = e^{-2}.$$

Q: Find  $P(3 < X \leq 8)$ .

A:

$$\begin{aligned} P(3 < X \leq 8) &= F_X(8) - F_X(3) \\ &= (1 - e^{-0.2 \cdot 8}) - (1 - e^{-0.2 \cdot 3}) \\ &= e^{-0.6} - e^{-1.6}. \end{aligned}$$

Q: A continuous random variable  $X$  has CDF

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-x}, & x \geq 0. \end{cases}$$

Find the PDF  $f_X(x)$ .

Q: A continuous random variable  $X$  has CDF

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-x}, & x \geq 0. \end{cases}$$

Find the PDF  $f_X(x)$ .

A: For a continuous RV,  $f_X(x) = F'_X(x)$  where it is differentiable.

So

$$f_X(x) = \begin{cases} 0, & x < 0, \\ \frac{d}{dx}(1 - e^{-x}) = e^{-x}, & x \geq 0. \end{cases}$$

Indeed  $\int_0^{\infty} e^{-x} dx = 1$ , so this is a valid PDF.

Q: A coffee shop gives a free drink coupon with probability  $\frac{1}{3}$ . If you *don't* get a coupon, then the amount you spend (in dollars) is equally likely anywhere between \$0 and \$6.

Let  $X$  be the amount you spend. Find  $P(X \leq 2)$ .

Q: A coffee shop gives a free drink coupon with probability  $\frac{1}{3}$ . If you *don't* get a coupon, then the amount you spend (in dollars) is equally likely anywhere between \$0 and \$6.

Let  $X$  be the amount you spend. Find  $P(X \leq 2)$ .

A: This is a mixed RV: there is a point mass at  $X = 0$  (coupon), plus a continuous part.

$$P(X \leq 2) = P(\text{coupon}) + P(0 < X \leq 2).$$

If no coupon,  $X \sim \text{Unif}(0, 6)$ , so

$$P(0 < X \leq 2 \mid \text{no coupon}) = \frac{2}{6} = \frac{1}{3}.$$

Also  $P(\text{no coupon}) = \frac{2}{3}$ , so

$$P(0 < X \leq 2) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}.$$

$$P(X \leq 2) = \frac{1}{3} + \frac{2}{9} = \frac{5}{9}.$$

Q: Let  $X \sim N(70, 9)$  (so  $\sigma = 3$ ). Find  $P(67 \leq X \leq 76)$  in terms of the standard normal CDF  $\Phi$ .

Q: Let  $X \sim N(70, 9)$  (so  $\sigma = 3$ ). Find  $P(67 \leq X \leq 76)$  in terms of the standard normal CDF  $\Phi$ .

A: Standardize:

$$P(67 \leq X \leq 76) = P\left(\frac{67 - 70}{3} \leq Z \leq \frac{76 - 70}{3}\right) = P(-1 \leq Z \leq 2).$$

So

$$P(-1 \leq Z \leq 2) = \Phi(2) - \Phi(-1).$$

- ▶ Know the discrete and continuous exemplary distributions
- ▶ Know different scenarios which they emulate
  - ▶ Often, these are tricky!
- ▶ Know the parameters that prescribe each distribution
- ▶ Know how to use the PDF and CDF, and convert between
- ▶ <https://www.acsu.buffalo.edu/~adamcunn/probability/probability.html>

## Expectation

Expectation describes the probabilistic “mean” of a RV.

$$E[X] = \sum_x x p_X(x). \quad (\text{disc.})$$

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx. \quad (\text{cont.})$$

The scaling/shift relationship is given by

$$E[aX + b] = aE[X] + b$$

Expectation is linear.

$$E[X + Y] = E[X] + E[Y].$$

### Variance

Variance is a measure of dispersion given by

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

The scaling/shift relationship is given by

$$\text{Var}(aX + b) = a^2 \text{Var}(X).$$

### Covariance

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y).$$

$\text{Var}(X) = 0$  if and only if  $X$  is degenerate (it only has 1 value).

Random variables  $X, Y$  are independent if **for all**  $x, y$ ,<sup>2</sup>

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y).$$

### Consequences

- ▶ Product of expectation holds.  $E[XY] = E[X]E[Y]$
- ▶ Covariance is zero.  $\text{Cov}(X, Y) = 0$

---

<sup>2</sup>We can't use  $P(X = x)$  as this is a general definition for discrete and continuous RVs

Distribution	Support	$E[X]$	$\text{Var}(X)$
Bernoulli( $p$ )	$\{0, 1\}$	$p$	$p(1 - p)$
Binomial( $n, p$ )	$\{0, 1, \dots, n\}$	$np$	$np(1 - p)$
Poisson( $\lambda$ )	$\{0, 1, \dots, \infty\}$	$\lambda$	$\lambda$
Geometric( $p$ )	$\{1, 2, \dots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Uniform( $a, b$ )	$[a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential( $\lambda$ )	$[0, \infty)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gaussian $N(\mu, \sigma^2)$	$\mathbb{R}$	$\mu$	$\sigma^2$

Q: Given  $E[X] = 2$ ,  $\text{Var}(X) = 3$ , what is  $E[X^2]$ ?

Q: The continuous random variable  $Y$  has PDF  $f(z) = 3z^2/19$  for  $[2, 3]$  and  $f(z) = 0$  otherwise. Find  $E[Y]$ .

Q: Given  $E[X] = 2$ ,  $\text{Var}(X) = 3$ , what is  $E[X^2]$ ?

A: Recall  $\text{Var}(X) = E[X^2] - E[X]^2$ , ergo  $3 + 2^2 = 7$

Q: The continuous random variable  $Y$  has PDF  $f(z) = 3z^2/19$  for  $[2, 3]$  and  $f(z) = 0$  otherwise. Find  $E[Y]$ .

A: Note that

$$E[Y] = \int_{\mathbb{R}} zf(z)dz = \int_2^3 z \frac{3z^2}{19} dz = \left[ \frac{3}{76} z^4 \right]_2^3 = \frac{195}{76}$$

Q: Ten people put their hats in a box and each selects one at random. Let  $X$  describe the number of people who get their hat back. What is  $E[X]$  and  $\text{Var}[X]$ ?

Q: Ten people put their hats in a box and each selects one at random. Let  $X$  describe the number of people who get their hat back. What is  $E[X]$  and  $\text{Var}[X]$ ?

A: We have 10 indicator random variables that person  $i$  gets their own hat back  $X_i \sim \text{Bern}(\frac{1}{n})$ . Note  $E[X_i] = \frac{1}{n}$

$$E[X] = \sum_i E[X_i] = \frac{n}{n} = 1$$

$$\begin{aligned} E[X^2] &= E \left[ \left( \sum_i X_i \right)^2 \right] = \sum_i E[X_i^2] + \sum_{i \neq j} E[X_i X_j] \\ &= n \left( \frac{1}{n} \right) + n(n-1) \left( \frac{1}{n(n-1)} \right) = 2 \end{aligned}$$

$$\text{Var}(X) = 2 - 1^2 = 1$$

**Calc Review: Integration by Parts**

$$\int u dv = uv - \int v du$$

Why? Product rule, the integrate

$$\frac{d}{dx}[uv] = uv' + u'v$$

$$\int \frac{d}{dx} uv dx = \int uv' dx + \int u' v dx$$

$$uv = \int uv' dx + \int u' v dx$$

Q: Let  $X \sim N(0, 1)$ . Show that  $E[X^2] = 1$  using integration by parts.

Q: Let  $X \sim N(0, 1)$ . Show that  $E[X^2] = 1$  using integration by parts.

A: Write the integral and integrate by parts.

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Use that  $\frac{d}{dx}(e^{-x^2/2}) = -xe^{-x^2/2}$ . Then

$$\begin{aligned} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx &= \int_{-\infty}^{\infty} x(xe^{-x^2/2}) dx = - \int_{-\infty}^{\infty} x d(e^{-x^2/2}) \\ &= - \left[ xe^{-x^2/2} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2} dx \\ &= \int_{-\infty}^{\infty} e^{-x^2/2} dx. \end{aligned}$$

Divide both sides by  $\sqrt{2\pi}$  to get

$$E[X^2] = \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx}_{\text{this is the normalization of } N(0, 1)} = 1.$$

Update: This is not on the exam.

### Moment

The  $k$ th **moment** of a random variable is given by

$$E[X^k].$$

The **moment generating function** is defined by

$$M_X(t) = E[e^{tX}], \quad \text{and if it exists near } 0, \quad \frac{d^k M_X(0)}{dt^k} = E[X^k].$$

If  $X$  is symmetric around 0, all odd moments are equal to 0.