Review for final exam : probability unit
MIT 18.05 Spring 2022

Sets and counting

- Sets: $\emptyset$, union, intersection, complement Venn diagrams, products
- Counting: inclusion-exclusion, rule of product, permutations ${}_nP_k$, combinations ${}_nC_k = \binom{n}{k}$

Problem 1. Consider the nucleotides A, G, C, T.

(a) How many ways are there to make a sequence of 5 nucleotides.

(b) How many sequences of length 5 are there where no adjacent nucleotides are the same

(c) How many sequences of length 5 have exactly one A?

Problem 2. **(a)** How many 5 card poker hands are there?

(b) How many ways are there to get a full house (3 of one rank and 2 of another)?

(c) What's the probability of getting a full house?

Problem 3. **(Counting)**
(a) How many arrangements of the letters in the word probability are there?

(b) Suppose all of these arrangements are written in a list and one is chosen at random. What is the probability it begins with 'b'.

Probability

- Sample space, outcome, event, probability function. Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
  Special case: $P(A^c) = 1 - P(A)$
  (A and B disjoint $\Rightarrow P(A \cup B) = P(A) + P(B)$.)
- Conditional probability, multiplication rule, trees, law of total probability, independence
- Bayes' theorem, base rate fallacy

Problem 4. Let E and F be two events. Suppose the probability that at least one of them occurs is 2/3. What is the probability that neither E nor F occurs?

Problem 5. Let C and D be two events with $P(C) = 0.3$, $P(D) = 0.4$, and $P(C^c \cap D) = 0.2$.

What is $P(C \cap D)$?

Problem 6. Suppose we have 8 teams labeled T_1, ..., T_8. Suppose they are ordered by placing their names in a hat and drawing the names out one at a time.

(a) How many ways can it happen that all the odd numbered teams are in the odd numbered slots and all the even numbered teams are in the even numbered slots?

(b) What is the probability of this happening?

Problem 7. More cards! Suppose you want to divide a 52 card deck into four hands with 13 cards each. What is the probability that each hand has a king?

Problem 8. Suppose we roll a fair die twice. Let A be the event 'the sum of the rolls is 5' and let B be the event 'at least one of the rolls is 4.'

(a) Calculate $P(A|B)$.

(b) Are A and B independent?

Problem 9. On a quiz show the contestant is given a multiple choice question with 4 options. Suppose there is a 70% chance the contestant actually knows the answer. If they don't know the answer they guess with a 25% chance of getting it right. Suppose they get it right. What is the probability that they were guessing?

Problem 10. Suppose you have an urn containing 7 red and 3 blue balls. You draw three balls at random. On each draw, if the ball is red you set it aside and if the ball is blue you put it back in the urn. What is the probability that the third draw is blue?

(If you get a blue ball it counts as a draw even though you put it back in the urn.)

Problem 11. Suppose that $P(A) = 0.4, P(B) = 0.3$ and $P((A \cup B)^C) = 0.42$. Are A and B independent?

Problem 12. Suppose now that events A, B and C are *mutually independent* with

$$P(A) = 0.3, \quad P(B) = 0.4, \quad P(C) = 0.5.$$

Compute the following: (Hint: Use a Venn diagram)
(i) $P(A \cap B \cap C^c)$ (ii) $P(A \cap B^c \cap C)$ (iii) $P(A^c \cap B \cap C)$

Problem 13. Suppose A and B are events with $0 < P(A) < 1$ and $0 < P(B) < 1$.

(a) If A and B are disjoint can they be independent?

(b) If A and B are independent can they be disjoint?

(c) If $A \subset B$ can they be independent?

Random variables, expectation and variance

- Discrete random variables: events, pmf, cdf

- Bernoulli(p), binomial(n, p), geometric(p), uniform(n)
- $E[X]$, meaning, algebraic properties, $E[h(X)]$
- Var(X), meaning, algebraic properties
- Continuous random variables: pdf, cdf
- uniform(a,b), exponential(λ), normal(μ,σ)
- Transforming random variables
- Quantiles

Problem 14. Directly from the definitions of expected value and variance, compute $E[X]$ and Var(X) when X has probability mass function given by the following table:

X	-2	-1	0	1	2
pmf	1/15	2/15	3/15	4/15	5/15

Problem 15. Suppose that X takes values between 0 and 1 and has probability density function $2x$. Compute Var(X) and Var(X^2).

Problem 16. The pmf of X is given by the following table

Value of X	-1	0	1
Probability	1/3	1/6	1/2

(a) Compute $E[X]$.

(b) Give the pdf of $Y = X^2$ and use it to compute $E[Y]$.

(c) Instead, compute $E[X^2]$ directly from an extended table.

(d) Compute Var(X).

Problem 17. Compute the expectation and variance of a Bernoulli(p) random variable.

Problem 18. Suppose 100 people all toss a hat into a box and then proceed to randomly pick out of a hat. What is the expected number of people to get their own hat back.

Hint: express the number of people who get their own hat as a sum of random variables whose expected value is easy to compute.

pmf, pdf, cdf

Probability Mass Functions, Probability Density Functions and Cumulative Distribution Functions

Problem 19. Suppose that $X \sim \text{Bin}(n, 0.5)$. Find the probability mass function of $Y = 2X$.

Problem 20. **(a)** Suppose that X is uniform on $[0,1]$. Compute the pdf and cdf of X.

(b) If $Y = 2X + 5$, compute the pdf and cdf of Y.

Problem 21. **(a)** Suppose that X has probability density function $f_X(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. Compute the cdf, $F_X(x)$.

(b) If $Y = X^2$, compute the pdf and cdf of Y.

Problem 22. Suppose that X is a random variable that takes on values 0, 2 and 3 with probabilities 0.3, 0.1, 0.6 respectively. Let $Y = 3(X-1)^2$.

(a) What is the expectation of X?

(b) What is the variance of X?

(c) What is the expection of Y?

(d) Let $F_Y(t)$ be the cumulative density function of Y. What is $F_Y(7)$?

Problem 23. Suppose you roll a fair 6-sided die 25 times (independently), and you get \$3 every time you roll a 6. Let X be the total number of dollars you win.

(a) What is the pmf of X.

(b) Find $E[X]$ and $\text{Var}(X)$.

(c) Let Y be the total won on another 25 independent rolls. Compute and compare $E[X+Y]$, $E[2X]$, $\text{Var}(X+Y)$, $\text{Var}(2X)$.

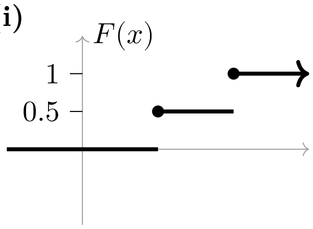
Explain briefly why this makes sense.

Problem 24. A continuous random variable X has PDF $f(x) = x + ax^2$ on [0,1]

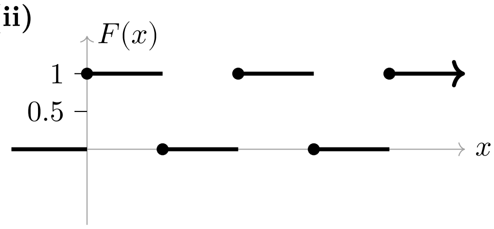
Find a, the CDF and $P(0.5 < X < 1)$.

Problem 25. For each of the following say whether it can be the graph of a cdf. If it can be, say whether the variable is discrete or continuous.

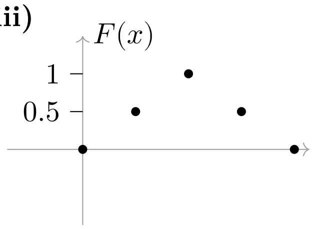
(i)



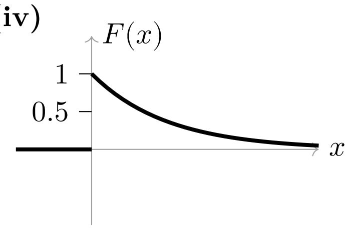
(ii)



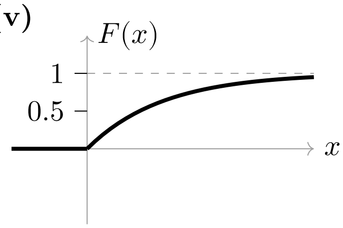
(iii)



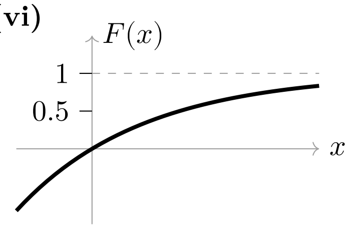
(iv)



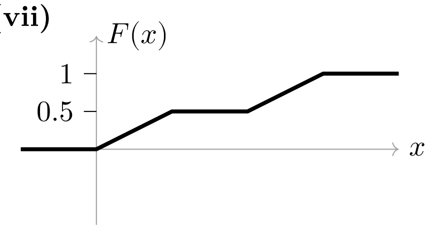
(v)



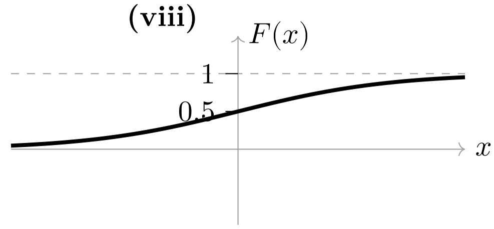
(vi)



(vii)



(viii)



Problem 26. Correlation
Flip a coin 5 times. Use properties of covariance to compute the covariance and correlation between the number of heads on the first 3 and last 3 flips.

Distributions with names

Problem 27. Exponential Distribution
Suppose that buses arrive are scheduled to arrive at a bus stop at noon but are always X minutes late, where X is an exponential random variable with probability density function $f_X(x) = \lambda e^{-\lambda x}$. Suppose that you arrive at the bus stop precisely at noon.

(a) Compute the probability that you have to wait for more than five minutes for the bus to arrive.

(b) Suppose that you have already waiting for 10 minutes. Compute the probability that you have to wait an additional five minutes or more.

Problem 28. Normal Distribution: Throughout these problems, let ϕ and Φ be the pdf and cdf, respectively, of the standard normal distribution Suppose Z is a standard normal random variable and let $X = 3Z + 1$.

(a) Express $P(X \le x)$ in terms of Φ

(b) Differentiate the expression from (a) with respect to x to get the pdf of X, $f(x)$. Remember that $\Phi'(z) = \phi(z)$ and don't forget the chain rule

(c) Find $P(-1 \le X \le 1)$

(d) Recall that the probability that Z is within one standard deviation of its mean is approximately 68%. What is the probability that X is within one standard deviation of its mean?

Problem 29. Transforming Normal Distributions
Suppose $Z \sim \mathrm{N}(0{,}1)$ and $Y = e^Z$.

(a) Find the cdf $F_Y(a)$ and pdf $f_Y(y)$ for Y. (For the CDF, the best you can do is write it in terms of Φ the standard normal cdf.)

(b) We don't have a formula for $\Phi(z)$ so we don't have a formula for quantiles. So we have to write quantiles *in terms* of Φ^{-1}.

(i) Write the 0.33 quantile of Z in terms of Φ^{-1}

(ii) Write the 0.9 quantile of Y in terms of Φ^{-1}.

(iii) Find the median of Y.

Problem 30. (**Random variables derived from normal random variables**)
Let X_1, X_2, ...X_n be i.i.d. N$(0, 1)$ random variables.

Let $Y_n = X_1^2 + ... + X_n^2$.

(a) Use the formula $\text{Var}(X_j) = E[X_j^2] - E[X_j]^2$ to show $E[X_j^2] = 1$.

(b) Set up an integral in x for computing $E[X_j^4]$.

For 3 extra credit points, use integration by parts show $E[X_j^4] = 3$.

(If you don't do this, you can still use this result in part c.)

(c) Deduce from parts (a) and (b) that $\text{Var}(X_j^2) = 2$.

(d) Use the Central Limit Theorem to approximate $P(Y_{100} > 110)$.

Problem 31. More Transforming Normal Distributions
(a) Suppose Z is a standard normal random variable and let $Y = aZ + b$, where $a > 0$ and b are constants.

Show $Y \sim \text{N}(b, a^2)$ (remember our notation for normal distributions uses mean and variance).

(b) Suppose $Y \sim \text{N}(\mu, \sigma^2)$. Show $\dfrac{Y - \mu}{\sigma}$ follows a standard normal distribution.

Problem 32. (**Sums of normal random variables**)
Let X, Y be independent random variables where $X \sim N(2, 5)$ and $Y \sim N(5, 9)$ (we use the notation $N(\mu, \sigma^2)$). Let $W = 3X - 2Y + 1$.

(a) Compute $E[W]$ and $\text{Var}(W)$.

(b) It is known that the sum of independent normal distributions is normal. Compute $P(W \leq 6)$.

Problem 33. Let $X \sim \text{U}(a, b)$. Compute $E[X]$ and $\text{Var}(X)$.