1. **(12 points)** You randomly shuffle a standard deck of 52 cards and you take the top 10 cards, in order, one by one. What is the probability that:

(a) The first card is a jack?

$$\frac{4}{52} = \frac{1}{13}$$

(b) The first card is **not** a diamond?

$$1 - \frac{13}{52} = \frac{39}{52} = \frac{3}{4}$$

(c) The first and second cards are the same suit? X

$$\frac{52C1}{52} \cdot \frac{12}{51} = \frac{12}{51}$$

(d) The fifth card is black?

$$\frac{26C1}{52} = \frac{1}{2}$$

(e) The first card is a spade, given that it is an ace?

$$P(\spadesuit \mid Ace) = \frac{P(\spadesuit \,\&\, Ace)}{P(Ace)} = \frac{\frac{1}{52}}{\frac{4}{52}} = \frac{1}{4}$$

(f) The second card is a queen, given that the first card is a queen?

$$\frac{3}{51} = \frac{1}{17}$$

2. **(3 points)** Ten cards are chosen at random from a standard deck. Which of the following pairs of events E_1, E_2 are independent? Circle them.

(a) E_1: first card is a 7, E_2: tenth card is a spade

- Independent

(b) E_1: second card is a club, E_2: fourth card is a heart

- Dependent

(c) E_1: first card is a heart, E_2: second card is a 10

- Independent

3. **(8 points)** Short answer questions.

(a) The letters C, D, U, S are randomly permuted. What is the probability that the result is $UCSD$?

$$\frac{1}{4!}$$

(b) Four fair dice are rolled. What is the probability that they all show the same number?

$$\frac{6}{6^4} = \frac{1}{6^3}$$

(c) Each time you go to the store, you have a 5% chance of running into your worst enemy. What is the expected number of trips to the store before you meet this person?

$$E(X): \frac{1}{p} = \frac{1}{0.05} = 20$$

(d) A certain population consists of 40% men and 60% women. Of the men, 30% are left-handed, and of the women, 10% are left-handed. A person is picked at random from this population and is found to be left-handed. What is the probability that this person is female?

Men $\to$ M, Women $\to$ W, Left handed $\to$ LH.

Let's calculate $P(LH) = P(LH \,\&\, M) + P(LH \,\&\, W) = P(m) \cdot P(LH|m) + P(w) \cdot P(LH|w)$

$= 0.4 \times 0.3 + 0.6 \times 0.1$

$= 0.18$

$\therefore \; P(w|LH) = P(w) \cdot \frac{P(LH|w)}{P(LH)} = 0.6 \times \frac{0.1}{0.18} = \frac{1}{3} \Rightarrow \boxed{P(w|LH) = \frac{1}{3} = 0.33}$

4. **(4 points)** You wake up one morning and find a horrible rash on your neck. You wonder, is it an insect bite? Here are the relevant facts:

- If you got an insect bite, you would definitely get a rash. $P(\text{rash}|\text{bite}) = 1$
- If you didn't get an insect bite, the probability of getting a rash (from other causes) is 1/8. $P(\text{rash}|\text{no bite}) = \frac{1}{8}$
- On any given night, in your house, the probability of getting an insect bite is 1/5. $P(\text{bite}) = \frac{1}{5}$

(a) On any given night, what is the probability of getting a rash?

$p(\text{rash}) = p(\text{bite} \,\&\, \text{rash}) + p(\text{no bite} \,\&\, \text{rash})$

$= p(\text{bite})\, p(\text{rash}|\text{bite}) + p(\text{no bite})\, p(\text{rash}|\text{no bite})$

$= \frac{1}{5} \times 1 + \frac{4}{5} \times \frac{1}{8} = \frac{3}{10} \Rightarrow \boxed{P(\text{rash}) = \frac{3}{10}}$

(b) Given that you got a rash, what is the probability that you got an insect bite?

$p(\text{insect bite}|\text{rash}) = p(\text{bite}) \cdot \frac{p(\text{rash}|\text{bite})}{p(\text{rash})}$

$= \frac{1}{5} \cdot \frac{1}{\frac{3}{10}}$

$\boxed{p(\text{insect bite}|\text{rash}) = \frac{2}{3}}$

5. **(8 points)** A die has six sides that come up with different probabilities.

$$\Pr(1) = \Pr(2) = \Pr(3) = \frac{1}{6}, \quad \Pr(4) = \Pr(5) = \frac{1}{8}, \quad \Pr(6) = \frac{1}{4}.$$

(a) You roll the die; let X denote the outcome. What is $\mathbb{E}(X)$?

$$E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{8} + 5 \times \frac{1}{8} + 6 \times \frac{1}{4} = \frac{87}{24}$$

$$\boxed{E(X) = 3.625}$$

(b) What is $\text{var}(X)$?

$$E(X^2) = 1 + \frac{1}{6} + 4 \times \frac{1}{6} + 9 \times \frac{1}{6} + 16 \times \frac{1}{8} + 25 \times \frac{1}{8} + 36 \times \frac{1}{4} = 16.458$$

$$\text{Var}(X) = E(X^2) - E(X)^2 = 16.458 - 3.625^2 = 3.31$$

$$\boxed{\text{var}(X) = 3.31}$$

(c) Now you roll this die a hundred times, and let Z be the **average** of all the rolls. What is $\mathbb{E}(Z)$?

$$E(Z) = \frac{100\, E(X)}{100}$$ according to linearity of expectation

$$\boxed{E(Z) = 3.625}$$

(d) What is $\text{var}(Z)$? since they are independent

$$\text{var}(Z) = \frac{100 \times \text{Var}(X)}{100}$$

$$\boxed{\text{Var}(Z) = 3.31}$$

6. **(4 points)** Recall that $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2. Random variables X, Y, Z are chosen independently as follows:

- X is chosen from $N(1, 16)$
- Y is chosen from $N(0, 4)$
- Z is chosen from $N(2, 9)$

Define $W = X + 2Y + Z$.

(a) What is $\mathbb{E}(W)$?

$$E(W) = E(X) + 2E(Y) + E(Z)$$
$$= 1 + 2(0) + 2$$

$$\boxed{E(W) = 3}$$

(b) What is $\text{var}(W)$?

$$\text{var}(W) = \text{var}(X) + 2^2 \cdot \text{var}(Y) + \text{var}(Z)$$
$$= 16 + 4 \times 4 + 9$$

$$\boxed{\text{Var}(W) = 41}$$

7. **(3 points)** A pair of random variables X_1 and X_2 have the following properties:

- They both take values in $\{-1, 1\}$
- X_1 has mean 0 while X_2 has mean 0.6
- The correlation between X_1 and X_2 is -0.25

Suppose we fit a (bivariate) Gaussian to (X_1, X_2). Give the mean and covariance matrix of this Gaussian.

$\mu = \begin{pmatrix} 0 \\ 0.6 \end{pmatrix}$

$Var(x_1) = 1 - 0$
$= 1$

$Var(x_2) = 1 - 0.6^2 = 0.64$

$\Sigma = \begin{pmatrix} 1 & -0.2 \\ -0.2 & 0.64 \end{pmatrix}$

$cov(x_1, x_2) = Corr(x_1, x_2) \cdot std(x_1) \cdot std(x_2)$
$= -0.25 \times 1 \times 0.8$
$= -0.2$

8. **(6 points)** Problems about data distributions.

(a) An auto-repair shop decides to keep track of the number of customers it receives every hour. During a particular stretch of 10 hours, these numbers (customers per hour) are:

$$3, 3, 1, 0, 0, 1, 2, 4, 0, 4.$$

0 1 2 3 4 5 6 7 8 9

They decide to fit a Poisson(λ) distribution to this data. What is the maximum-likelihood λ?

λ = Average of the values. (impirical mean)
= 1.8

(b) Random vector $X = (X_1, X_2)$ is drawn according to the bivariate Gaussian with the following mean and covariance matrix:

$$\mu = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$$

What is the probability that $X_2 \geq 4$?

$P(X_2 \geq 4)$

Since 4 is mean (μ) for X_2. $P(X_2 \geq 4) = \frac{1}{2} = 0.5$

(c) Fix vocabulary V = {`beside, dog, cat, another`}. Using Laplace smoothing, what multinomial distribution would you fit to the following sentence?

`Beside the dog were a cat, another dog, and yet another dog.`

Laplace smoothing.

	Beside	the	dog	were	a	cat	another	and	yet
# of occurance	1	1	3	1	1	1	2	1	1
Pr.	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
Pr(word) Laplace smoothing $\frac{\#occur + 1}{\text{Total words} + \lvert V \rvert}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{2}{16}$

9. **(2 points)** A school wants to determine the average number of hours that the students spend on homework; call this unknown number μ. 100 students are chosen at random, and each of them is asked to report the typical number of hours per week that he or she spends on homework. The reported numbers have a mean of 10.1 and a standard deviation of 6.2. Give a 95% confidence interval for μ.

$n = 100 \qquad \mu = 10.1 \qquad \text{stdd} = 6.2 \qquad \sigma = 6.2.$

for 95% confidence interval.

$\mu \pm 2\sigma, \quad [10.1 - 2\times 6.2, \ 10.1 + 2\times 6.2]$

$= [-2.3, \ 22.5]$

10. **(6 points)** Consider the following two unit vectors in $\mathbb{R}^3$:

$$u_1 = \begin{pmatrix} 3/5 \\ 4/5 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} -4/5 \\ 3/5 \\ 0 \end{pmatrix}.$$

(a) Notice u_1 and u_2 are orthogonal. Find a third unit vector u_3 that is orthogonal to both u_1 and u_2.

$$u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

(b) A particular data set in $\mathbb{R}^3$ has covariance matrix M with eigenvectors u_1, u_2, u_3 (from part (a)) and eigenvalues 4, 2, 1, respectively. Suppose principal component analysis is used to project this data set into $\mathbb{R}^2$. What is the 2-dimensional projection of $x = (10, 5, 5)$?

Projection into 2d : $U^T x$

$$= \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \end{pmatrix} \begin{pmatrix} 10 \\ 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 10 \\ -5 \end{pmatrix}$$

(c) What is the three-dimensional reconstruction of x from the projection you obtained in part (b)?

reconstruction $= UU^T x$

$$= \begin{pmatrix} 10 \\ 5 \\ 0 \end{pmatrix}.$$

11. **(4 points)** A survey is taken to determine what fraction of freshman computer science majors have prior programming experience. Call this unknown fraction p. Out of the nationwide pool of computer science freshmen, 100 are chosen at random. Of them, 36% had prior programming experience.

(a) The natural estimate of p is 0.36. Give a 95% confidence interval for the estimate.

$P = 0.36$, $\mu = p$
$= 0.36$

$$\sigma = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.36 \times 0.64}{100}} = 0.048$$

Confidence interval for 95% $= \mu \pm 2\sigma$
$= [0.264, 0.456]$

(b) Suppose we now want to estimate p more accurately, to within a 95% confidence interval of ± 0.05. What sample size should we use?

for chance error 0.05

$$2\sigma \leq 0.05 \Rightarrow 2\sqrt{\frac{0.36 \times 0.64}{n}} \leq 0.05$$

$n \geq 368.64 \approx \boxed{n \geq 369}$

12. **(20 points)** For this last problem, you should turn in an iPython notebook.

Download the IRIS data set from:

`https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data`

This is a data set of 150 points in $\mathbb{R}^4$, with three classes; for more details of the features and classes, refer to the website `https://archive.ics.uci.edu/ml/datasets/iris`.

Build a classifier for this data set, based on a generative model (you can choose whichever you like).

- Split the data set into training/test data as follows: use the first 35 points in each class for training, and use the remaining 15 points for testing.
- What error rate do you achieve?