Problems for Griffiths' Electrodynamics

Contents

481 Lecture 1/16/24

1 Chapter 1: Probablities and Interference (Mackay Ch 2-3)

An ensemble: x random variable

$$A_x = (a_1, a_2, \dots, a_n)$$
$$P_x = (p_1, p_2, \dots, p_n)$$
$$p(x = a_i) = p_i$$

x takes value a_i with probability p_i

$$p \geq 0, \quad \sum_{a_i \in A_x} p(x = a_i) = 1$$

Short hand for $p(x = a_i)$ is $p(a_i), \quad p(x)$
Joint ensemble: X, Y ensembles

$$XY = \text{ordered pairs}(x, y) \quad x \in A_X, y \in A_Y$$
$$P(x, y) = \text{joint probability of x and y}$$

Marginal probability: $P(x, y) \to P(x), P(y)$

$$P(x) = \sum_{y \in A_y} P(x, y)$$
$$P_x(x = a_i) = \sum_{b \epsilon A_y} P_{XY}(x = a_i, y = b)$$

Conditional probability:

$$P(x = a_i | y = b_j) = \frac{P(x = a_i, y = b_j)}{P(y = b_j)}$$

"Probability of $x = a_i$ given that $y = b_j$ (is true)"

Example 1 $XY = 2$ successive letters in english alphabet. P_x and P_y are identical 'frequency of a letter in english'

$$A_{xy} = \{aa, ab, ac, \dots, zz\}$$

$$P(y | x = `q')$$

Peak at $y = `u'$

$$\neq P_Y(y)$$

because x and y are not independent
X, Y "independent" if (and only if) $P(x, y) = P(x)P(y)$
Userful relations: $P(x, y) = P(x|y)P(y) = P(y|x)P(x)$
For any assumption H

$$\forall H : \qquad P(x, y|H) = p(x, y|H)p(y|H)$$

'Sum rule':

$$P(x|H) = \sum_{y \in A_y} P(x, y|H) = \sum_{y \in A_y} P(x|y, H)P(y|H)$$

2 Lecture 1/18

Last time: Main point $P(y|x) \neq P(y)$
Useful relations: Conditional probability

$$P(x|y) = \frac{P(x,y)}{P(y)}$$

where the joint relation is

$$P(x,y) = P(x|y)P(y) = P(y|x)P(x)$$

this can be rewritten into *Baye's theorem*

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

Example 2: Apply Baye's theorem Alex is test for a nast disease.

- Disease status: a (sick or healthy)
- Test outcome: b (positive or negative)

"Test is 95% reliable" or

$$P(+|\text{sick}) = 0.95, \quad P(-|\text{healthy}) = 0.95$$

Disease is nasty but rare $P(\text{sick}) = 0.01$; $P(\text{Healthy}) = 0.99$
Test is positive, what is the probability that Alex is sick? $P(\text{sick}|+) =?$

Solution Use Baye's theorem:

$$P(\text{sick}|+) = \frac{P(+|\text{sick})P(\text{sick})}{P(+)}$$

where $P(+)$ is the probability of a positive test result. This can be found using the sum rule

$$P(+) = P(+|\text{sick})P(\text{sick}) + P(+|\text{healthy})P(\text{healthy})$$

Thus

$$P(\text{sick}|+) = \frac{0.95 * 0.01}{0.95 * 0.01 + 0.05 * 0.99} = 0.161$$

It is useful to write the probabilities in a table

	$b = +$	$b = -$	$P(b)$
$a = \text{sick}$	$0.95 * 0.01$	$0.05 * 0.01$	0.01
$a = \text{healthy}$	$0.05 * 0.99$	$0.95 * 0.99$	0.99
$P(a)$	0.161	0.839	1

where columns represent the 95:5 reliable test.

Exclam!

$$P(S|+) \neq P(+|S)$$

A brief philosphical interlude. . . The 'Bayesian viewpoint':
Probability as degree of beliefs in propositions given assumptions & evidence, or Probability as 'freq of outcomes in repeat random experiments'

Forward and inverse problems

So far we have talked about Cond Prob, Baye's thrm, and and example.

Generative Model: Parameters $\Theta \to P(D|\Theta) \to (P)$ outcomes (data) AKA 'forward problem'
'a model' predicts an outcome given parameters. The model is a probablity distribution due to all the uncertainties and errors we have in the real world.

The Inverse Problem $P(\Theta|D)$
The inverse problem is the opposite of the forward problem (obviously). Also related to the issues regarding 'inference' and using Baye's theorem.

Example 3: A forward problem
An urn contains K balls, B balls are black, and $K - B$ balls are white. A ball is drawn at N times with replacement.

- $n_B = \#$ of times a black ball is drawn
- $P(n_B)$, average n_B?, STD?

With

$$f_B = \frac{B}{K}$$

The probability is given by the binomial distribution

$$P(n_B|N, f_B) = \binom{N}{n_B} f_B^{n_B}(1 - f_B)^{N-n_B}$$

The mean is $N * f_B$ and the STD is $\sqrt{N * f_B * (1 - f_B)}$

Example 4: An inverse problem
We have 11 urns, each with 10 balls. u is the number of black balls in each urn and the urns have $u = 0, 1, \ldots, 10$ black balls. Alex selects an urn at random and draws N balls at random with replacement. Bob wates Alex, but does not know which urn u was selected. For Bob, what is $P(u|N, n_B)$?
We have the data, but we are trying to infer the parameter u

Solution Use Baye's theorem

$$P(u|N, n_B) = \frac{P(n_B|u)P(u)}{P(n_B)}$$

where $P(n_B|u)$ is the 'forward' part from Ex 2, $P(u) = 1/11$, and $P(n_B)$ is the 'normalization' that makes it a valid prob. distribution:

$$P(n_B) = \sum_{u'} P(n_B|u')P(u')$$

Therefore

$$P(u|N, n_B) \propto \binom{N}{n_B}\left(\frac{u}{10}\right)^{n_B}\left(1 - \frac{u}{10}\right)^{N-n_B}$$

e.g. $n_B = 3, N = 10$
insert figure 1.2
The (0,0) point is impossible because we picked 3 black balls, and the urn $u = 0$ has no black balls. The same is true for the (10,10) point. The most likely point is $u = 3 \ldots$

Exclam! This is known as 'Posterior Probabilty'

- Θ is the parameter
- D is the data
- $P(\Theta)$ is the prior
- $P(D|\Theta)$ is the likelihood: a function of D prob of data given param (sums to 1 over all options for D). As a function of $\Theta \to$ likelihood of Θ
- $P(\Theta|D)$ is the posterior
- $P(D)$ is the normalization

! **Probability of *data***

! **Likelihood of *parameters***

Role of Prior:

! You can't do inference without making assumptions

Lecture 1/23/24

Last time:

- Forward $p(data|param)$
- Inverse $p(param|data)$

Using Baye's theorem

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} = \frac{\text{likelihood} \cdot \text{prior}}{\text{norm}}$$

Note: You can't do inference w/o working assumptions (prior) priors are subjective. From the inverse problem ex from last week: what is the probability that next ball Alex draws is black?

$$P(B) = \sum P(u)P(B|u)$$

Note: Infereince $\neq$ decision/choice of model. Inference is assigning probabilties to hypothesese.

Problem USB Cable frustrations "It takes 3 tries to plug in a USB cable"
During our first try to plug in the cable, we are collecting data. And if its wrong, we 'believe' that the orientation is wrong, thus we flip it believing that the 2nd try is the correct one. But in fact, this is wrong and the 3rd try is the correct one.

How to collect data?

Lecture 1/25/24

---

3 Chapter 2: Probabilities and Interference (Mackay Ch 2-3)

Example 5: Tossing a coin

- 3 times: H, H, H
- 10 times: H, H, ... H

what is the probability of the next toss being H?

Ex 5.1 Coin with freq of heads f_H is tossed N times and n_H heads. What is the probability of the next toss being H? (Ex 4 but with fixed unknown parameter)
Prior: subjective assumption (e.g. could be uniform) then do inference.

Ex 5.2 N tosses, n_H heads. What is the probability that the coin is biased? (Model Comparison)

Homework 1

Due 1/30 12pm

1. (a) For case 1.1 the marginal probabilty is

$$P(x=0)=0.2, \quad P(x=1)=0.8$$
$$P(y=0)=0.6, \quad P(y=1)=0.4$$

For 1.2

$$P(x=0)=0.4, \quad P(x=1)=0.6$$
$$P(y=0)=0.6, \quad P(y=1)=0.4$$

(b) 1.1

$$P(x=0|y=0)=1/5, \quad P(x=1|y=0)=4/5$$
$$P(x=0|y=1)=1/5, \quad P(x=1|y=1)=4/5$$

1.2

$$P(x=0|y=0)=1/3, \quad P(x=1|y=0)=2/3$$
$$P(x=0|y=1)=3/7, \quad P(x=1|y=1)=4/7$$

(c) Variables x and y are independent iff $P(x,y)=P(x)P(y)$. For 1.1

$$P(x=0,y=0)=0.12 \quad \text{and} \quad P(x=0)P(y=0)=0.2(0.6)=0.12$$
$$P(x=0,y=1)=P(x=0)P(y=1)=0.2(0.4)=0.08\ldots$$
$$P(x,y)=P(x)P(y)$$

So x and y are independent for 1.1. You can also see that the condition of y does not change the marginal probability of x. For 1.2 there is a simple counterexample

$$P(x=0,y=0)=0.1 \quad \text{and} \quad P(x=0)P(y=0)=0.4(0.3)=0.12$$
$$P(x=0,y=0)\neq P(x)P(y)$$

So x and y are not independent (dependent) for 1.2. You can also see that the conditional probability is not the same as the marginal probability for both cases.

2. For two random variables x and y to be independent, it must be true that

$$P(x,y)=P(x)P(y) \tag{1}$$

and from the definition of conditional probability

$$P(x|y=y_o)=\frac{P(x,y=y_o)}{P(y=y_o)}$$

substituting (1) into the joint probability

$$P(x|y=y_o)=\frac{P(x)P(y=y_o)}{P(y=y_o)}$$
$$P(x|y=y_o)=P(x)$$

3. (a) Since the two thrown dice are independent, the fair dice has 36 possible outcomes $A_{xy} = \{(1,1),(1,2),\ldots,(6,6)\}$ with equal probability

$$P(x,y) = P(x)P(y) = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$$

The probability distribution of the sum of the two dice $P(S)$ is

$$P(S) = \begin{cases} 1/36 & S = 2, 12 \\ 2/36 & S = 3, 11 \\ 3/36 & S = 4, 10 \\ 4/36 & S = 5, 9 \\ 5/36 & S = 6, 8 \\ 6/36 & S = 7 \end{cases}$$

where $S = x + y$. For the absolute difference of the two dice $D = |x - y|$

$$P(D) = \begin{cases} 2/36 & D = 5 \\ 4/36 & D = 4 \\ 6/36 & D = 3 \\ 8/36 & D = 2 \\ 10/36 & D = 1 \\ 6/36 & D = 0 \end{cases}$$

for the difference $D = 0$ there are 6 possible outcomes $(1,1),(2,2),\ldots,(6,6)$, for $D = 1$ there are 10 possible outcomes $(1,2),(2,1),(2,3),(3,2),\ldots,(5,6),(6,5)$, and so on.
(b) For 100 dice, the probability distribution of the sum of the dice $P(S)$ would be roughly

$$P(S) = \begin{cases} 1/6^{100} & S = 100, 600 \\ 100/6^{100} & S = 101, 599 \\ 5050/6^{100} & S = 102, 598 \\ \vdots & \vdots \\ 1.52\times 10^{76}/6^{100} & S = 350 \end{cases}$$

first we find the mean of 1 independent dice roll (μ_1)':

$$\mu_1 = \sum_x P(x)x = \frac{1}{6}\sum_{x=1}^{6} x = \frac{21}{6} = 3.5$$

because the mean of N independent dice rolls is the sum of the means of each dice roll

$$\mu_N = \sum_{i=1}^{N}\mu_i$$

thus the mean of 100 dice rolls is

$$\mu_{100} = 100\cdot\mu_1 = \boxed{350}$$

To find the Standard Deviation we first find the variance of 1 independent dice roll (σ_1^2):

$$\begin{aligned}[x] = [(x-[x])^2] &= [x^2 - 2x[x] + [x]^2] \\ &= [x^2] - 2[x]^2 + [x]^2[1] = [x^2] - [x]^2\end{aligned}$$

or in summation notation

$$\sigma_1^2 = \sum_x P(x)(x-\mu_1)^2 = \frac{1}{6}\sum_{x=1}^{6}(x-3.5)^2 = \frac{17.5}{6} = 2.9167$$

for 2 independent variables x and y

$$
\begin{aligned}
[x+y] &= [((x+y) - [x+y])^2] \\
&= [(x - [x]) + (y - [y])^2] \\
&= [(x - [x])^2 + (y - [y])^2 + 2(x - [x])(y - [y])] \\
&= [(x - [x])^2] + [(y - [y])^2] + 2[(x - [x])(y - [y])] \\
&= [x] + [y] + 2[(x - [x])(y - [y])]
\end{aligned}
$$

where the third term is

$$
\begin{aligned}
[(x - [x])(y - [y])] &= [xy - x[y] - y[x] + [x][y]] \\
&= [xy] - [x][y]
\end{aligned}
$$

and for independent variables x and y the third term is zero. Thus the variance of the sum of N independent dice rolls is

$$[N] = N\sigma_1^2 = 100 \cdot 2.9167 = 291.67$$

and the standard deviation is

$$\sigma_N = \sqrt{[N]} = \sqrt{291.67} = \boxed{17.08}$$

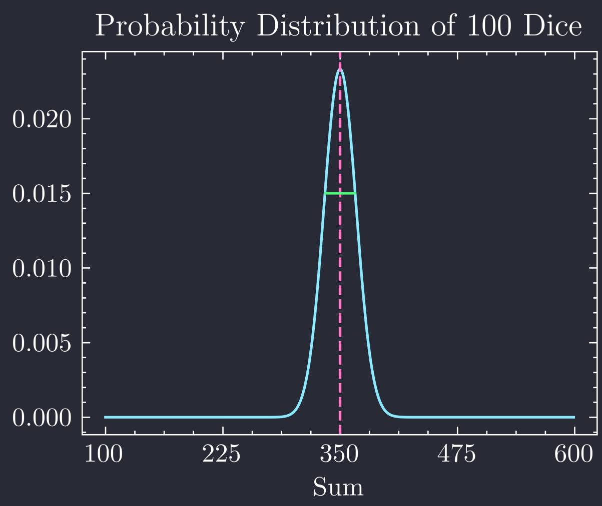

Figure 1.1: Probability distribution of the sum of 100 dice rolls. The mean is 350 and the standard deviation is ≈ 17.

The sketch of probability distribution of the sum of 100 dice rolls is shown in Figure 1.1.

4. (a) Assuming that there is an equal likelihood of the order of the age of the three brothers being any of the 6 possible permutations:

$$\text{Age}_{A,B,F} = \{ABF, AFB, BAF, BFA, FAB, FBA\}$$

where we denote the first element of a permutation as the oldest brother and the last element as the youngest brother.
The probability that Fred (F) is older than Bob (B) is $\boxed{1/2}$ from both the 3 possible permutations or by realizing that there are only two equal outcomes when looking at only the age of Fred vs Bob.
(b) Given that Fred is older than Alex (A), we can eliminate the 3 permutations where Alex is older. Thus the probability that Fred is older than Bob is $\boxed{2/3}$.

5. (a) Given that the probability of choosing a black ball from an urn is $f_B = \frac{B}{K}$, the probability distribution of choosing n_B black balls from N draws is

$$\boxed{P(n_B|N, f_B) = \binom{N}{n_B} f_B^{n_B}(1-f_B)^{N-n_B}}$$

(b) Finding the mean and standard deviation of n_B is quite similar to Problem 3. Since each draw is independent, the mean of n_B is the sum of the means of each draw! In the case of drawing one ball has a binary outcome

$$n_B(N=1) = \begin{cases} 1 & \text{black ball with probability } f_b \\ 0 & \text{white ball } (1-f_B) \end{cases}$$

thus the mean of n_B for one draw is

$$\mu_1 = 1(f_B) + 0(1-f_B) = f_B$$

and the variance is

$$\sigma_1^2 = (1-f_B)^2 f_B + (0-f_B)^2(1-f_B) = f_B(1-f_B)$$

for N draws the means add up to

$$\boxed{\mu = N f_B}$$

and the variances add only due to the independent nature of the draws

$$\sigma^2 = N f_B(1-f_B)$$

thus the standard deviation is

$$\boxed{\sigma = \sqrt{N f_B(1-f_B)}}$$

For $K = 20$, and $B = K$; $f_B = 5/20 = 0.25$. And for $N = 5$ we have the ratio

$$\frac{\sigma}{\mu} = \frac{\sqrt{5 \cdot 0.25 \cdot 0.75}}{5 \cdot 0.25} = \frac{\sqrt{15}}{5} \approx \boxed{0.77}$$

and for $N = 20$

$$\frac{\sigma}{\mu} = \frac{\sqrt{1000 \cdot 0.25 \cdot 0.75}}{1000 \cdot 0.25} = \frac{\sqrt{30}}{100} \approx \boxed{0.05}$$

6. (a) Dividing the time period T in to M intervals where each interval has a probability $r\,\mathrm{d}t$ or $\mathrm{d}t = T/M$. The probability of no events occuring in time T is

$$\lim_{M\to\infty} (1 - r\,\mathrm{d}t)^M = \lim_{M\to\infty} \left(1 - \frac{rT}{M}\right)^M = \boxed{e^{-rT}}$$

(b) For $n_T = x$ events occuring in M this is similar to the binomial distribution as $M \to \infty$.

$$\lim_{M\to\infty} \frac{M!}{x!(M-x)!} \left(\frac{rT}{M}\right)^x \left(1 - \frac{rT}{M}\right)^{M-x}$$

canceling out some terms...

$$\frac{M!}{(M-x)!} = \frac{M(M-1)\cdots(M-x+1)(M-x)!}{(M-x)!} = M(M-1)\cdots(M-x+1)$$

and now the facorial has x terms, so we can write it as

$$\frac{M(M-1)\cdots(M-x+1)}{M^x} = \frac{M}{M}\frac{M-1}{M}\cdots\frac{M-x+1}{M}$$
$$= 1\cdot\left(1-\frac{1}{M}\right)\cdots\left(1-\frac{x-1}{M}\right)$$

and as $M\to\infty$ the terms in the product go to 1, so the product goes to 1. Thus we are left with

$$\lim_{M\to\infty}\frac{(rT)^x}{x!}\left(1-\frac{rT}{M}\right)^{M-x} = \lim_{M\to\infty}\frac{(rT)^x}{x!}\left(1-\frac{rT}{M}\right)^{M}\left(1-\frac{rT}{M}\right)^{-x}$$

the second term is the limit of the exponential function

$$\lim_{M\to\infty}\left(1-\frac{rT}{M}\right)^{M} = e^{-rT}$$

and the third term tends to 1 as $M\to\infty$. Thus the probability of x events occuring in time T is

$$\boxed{P(x) = \frac{(rT)^x}{x!}e^{-rT}}$$

where $x = n_T$ for the sake of brevity in notation.
(c) The mean of n_T is

$$\begin{aligned}
\mu &= \sum_{x=0}^{\infty} xP(x) \\
&= \sum_{x=0}^{\infty} x\frac{(rT)^x}{x!}e^{-rT} \\
&= e^{-rT}\sum_{x=1}^{\infty} x\frac{(rT)(rT)^{x-1}}{x(x-1)!} \\
&= e^{-rT}(rT)\sum_{x=1}^{\infty}\frac{(rT)^{x-1}}{(x-1)!}
\end{aligned}$$

the the first term of the sum is zero which is why the sum starts at $x = 1$. The sum is also the Taylor series expansion of e^{rT} if we let $n = x-1$ so

$$\sum_{x=1}\frac{(rT)^{x-1}}{(x-1)!} = \sum_{n=0}\frac{(rT)^n}{n!} = e^{rT}$$

Therefore the mean of n_T is

$$\mu = (rT)e^{-rT}e^{rT} = rT$$

The variance of n_T is

$$\sigma^2 = [x^2] - [x]^2$$

the first term is solved similarly to the mean

$$\begin{aligned}
[x^2] &= \sum_{x=0}^{\infty} x^2 P(x) \\
&= \sum_{x=0}^{\infty} x^2 \frac{(rT)^x}{x!} e^{-rT} \\
&= e^{-rT} \sum_{x=1}^{\infty} x^2 \frac{(rT)(rT)^{x-1}}{x(x-1)!} \\
&= (rT)e^{-rT} \sum_{x=1}^{\infty} x \frac{(rT)^{x-1}}{(x-1)!} \\
&= (rT)e^{-rT} \left[\sum_{x=1}^{\infty} (x-1) \frac{(rT)^{x-1}}{(x-1)!} + \sum_{x=1}^{\infty} \frac{(rT)^{x-1}}{(x-1)!} \right] \qquad x = [(x-1)+1] \\
&= (rT)e^{-rT} \left[(rT) \sum_{x=2}^{\infty} \frac{(rT)^{x-2}}{(x-2)!} + \sum_{n=0}^{\infty} \frac{(rT)^n}{n!} \right] \qquad n = x-1 \\
&= (rT)e^{-rT} \left[(rT) \sum_{l=0}^{\infty} \frac{(rT)^l}{l!} + \sum_{n=0}^{\infty} \frac{(rT)^n}{n!} \right] \qquad l = x-2 \\
&= (rT)e^{-rT} \left[(rT)e^{rT} + e^{rT} \right] \\
&= (rT)^2 e^{-rT} e^{rT} + (rT)e^{-rT} e^{rT} \\
&= (rT)^2 + rT
\end{aligned}$$

and the variance is

$$\sigma^2 = (rT)^2 + rT - (rT)^2 = rT$$

Therefore the mean and standard deviation of n_T are

$$\boxed{\mu = rT \quad \text{and} \quad \sigma = \sqrt{rT}}$$

7. Using Bayes' theorem for the outcome $X = \{7, 3, 4, 2, 5, 3\}$ is

$$P(A|7,3,4,2,5,3) = P(A|X) = \frac{P(X|A)P(A)}{P(X)}$$

where the probability of choosing dice A is 1 in 3—$P(A) = 1/3$. The conditional probabilty $P(X|A)$ is the probability of rolling the outcome X given that dice A is chosen:

$$P(X|A) = \frac{1 \times 4 \times 2 \times 4 \times 2 \times 4}{20^6} = \frac{256}{20^6}$$

and the probability of rolling the outcome X is given by the sum rule

$$P(X) = P(X|A)P(A) + P(X|B)P(B) + P(X|C)P(C)$$

where $P(B) = P(C) = 1/3$ and the conditional probabilities $P(X|B)$ and $P(X|C)$ are

$$\begin{aligned}
P(X|B) &= \frac{2 \times 3 \times 2 \times 2 \times 2 \times 3}{20^6} = \frac{144}{20^6} \\
P(X|C) &= \frac{2^6}{20^6} = \frac{64}{20^6}
\end{aligned}$$

thus the probability of choosing dice A given the outcome X is

$$P(A|X) = \frac{\frac{256}{20^6} \cdot \frac{1}{3}}{\frac{256}{20^6} \cdot \frac{1}{3} + \frac{144}{20^6} \cdot \frac{1}{3} + \frac{64}{20^6} \cdot \frac{1}{3}} = \frac{256}{464} \approx \boxed{0.55}$$

with the knowledge that terms cancel out, the probability of the die being B is

$$P(B|X) = \frac{144}{464} \approx \boxed{0.31}$$

and the probability of the die being C is

$$P(C|X) = \frac{64}{464} \approx \boxed{0.14}$$

8. (a) Given that the bus arrives on average every 5 minutes, the average wait time is 5 minutes. And the bus that just left Sally would have left an average of 5 minutes ago. From the code, taking the mean value of the wait times is also $\approx$ 5 minutes.
(b) Therefore the average time between two buses is the sum in the time Sally is waiting for the bus and how long the missed bus has been gone: 10 minutes.
(c) The paradox is that 'we' think that after waiting for 5 minutes the bus will arrive, but the average time between buses is 10 minutes, so we are waiting longer than we expect to intuitively. This is because the conditional probability of Sally getting to the bus stop where the interval between buses is less than 5 minutes given that she has waited for a time t is less as time goes on. And the probability that Sally arrived at the bus stop where the interval between buses is more than 5 minutes given that she has waited for a time t is more as time goes on.
(d)

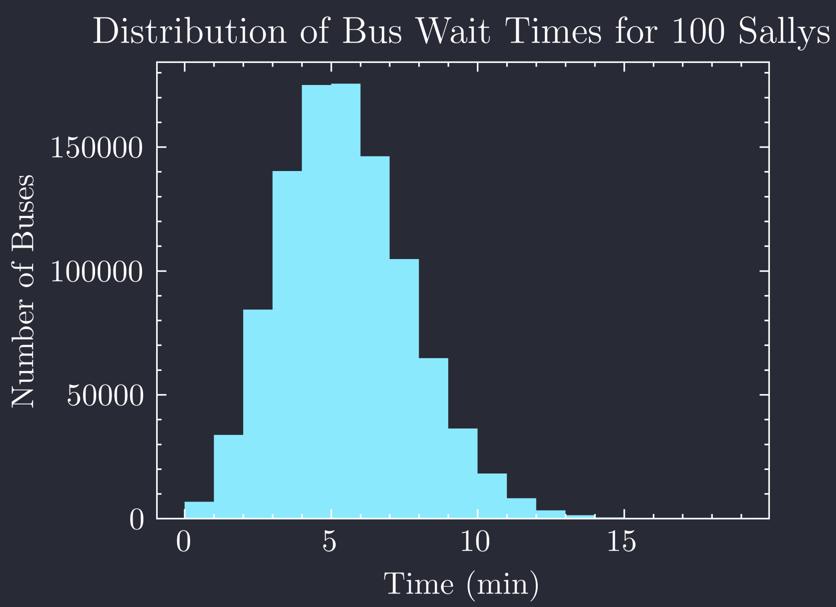

Figure 1.2: Mean of 5.00 min and Time between buses of 10.00 min.