Homework_11

April 15, 2025

Probability for Data Science

UC Berkeley, Spring 2025

Michael Xiao and Ani Adhikari

[1]:
```
import warnings
warnings.filterwarnings('ignore')

from prob140 import *
from datascience import *
import numpy as np
from scipy import stats

import matplotlib.pyplot as plt
%matplotlib inline
import matplotlib
matplotlib.style.use('fivethirtyeight')
```

1 Homework 11 (Due Monday, April 14th at 5 PM)

1.0.1 How to Do Your Homework

The point of homework is for you to try your hand at using what you've learned in class. The steps to follow:

- Go to lecture and sections, and also go over the relevant text sections before starting on the homework. This will remind you what was covered in class, and the text will typically contain examples not covered in lecture. The weekly Study Guide will list what you should read.
- Work on some of the practice problems before starting on the homework.
- Attempt the homework problems by yourself with the text, section work, and practice materials all at hand. Sometimes the week's lab will help as well. The two steps above will help this step go faster and be more fruitful.
- At this point, seek help if you need it. Don't ask how to do the problem — ask how to get started, or for a nudge to get you past where you are stuck. Always say what you have already tried. That helps us help you more effectively.

- For a good measure of your understanding, keep track of the fraction of the homework you can do by yourself or with minimal help. It's a better measure than your homework score, and only you can measure it.

1.0.2 Rules for Homework

- Every answer should contain a calculation or reasoning. For example, a calculation such as $(1/3)(0.8) + (2/3)(0.7)$ or `sum([(1/3)*0.8, (2/3)*0.7])` is fine without further explanation or simplification. If we want you to simplify, we'll ask you to. But just $\binom{5}{2}$ by itself is not fine; write "we want any 2 out of the 5 frogs and they can appear in any order" or whatever reasoning you used. Reasoning can be brief and abbreviated, e.g. "product rule" or "not mutually exclusive."
- You may consult others (see "How to Do Your Homework" above) but you must write up your own answers using your own words, notation, and sequence of steps.
- We'll be using Gradescope. You must submit the homework according to the instructions at the end of homework set.

1.1 We will not grade assignments which do not have pages correctly selected for each question.

1.1.1 Preliminary: Line Plots

In Data 8 you used `Table.plot` to draw line plots, but in this class the line plots have typically been drawn using the plot function of matplotlib. Here is the syntax; you are going to need it in the exercises. It's easier than setting up tables first, especially when you want to overlay multiple plots.

The `pyplot` module of matplotlib has been imported for you as `plt`. This is its standard abbreviation.

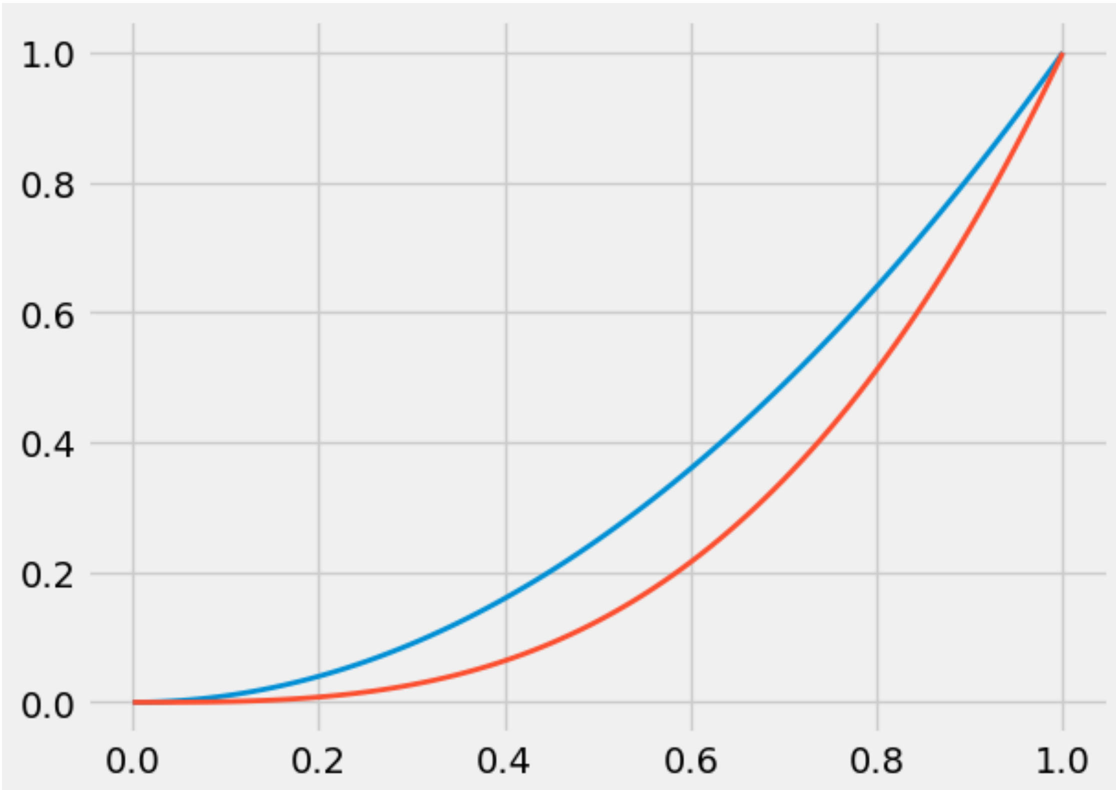
Let `x` and `y` be two numerical arrays of the same length. Then `plt.plot(x, y)` produces a line plot with values of `x` on the horizontal axis and the corresponding values of `y` on the vertical. The plot "joins the dots" `(x.item(0), y.item(0))`, `(x.item(1), y.item(1))`, `(x.item(2), y.item(2))`, and so on.

The optional argument `lw=n` can be used to set a line width of n units. Please use `lw=2` in this homework.

The semicolon at the end of the last line of code in each cell stops matplotlib from blurting out text that we don't need here.

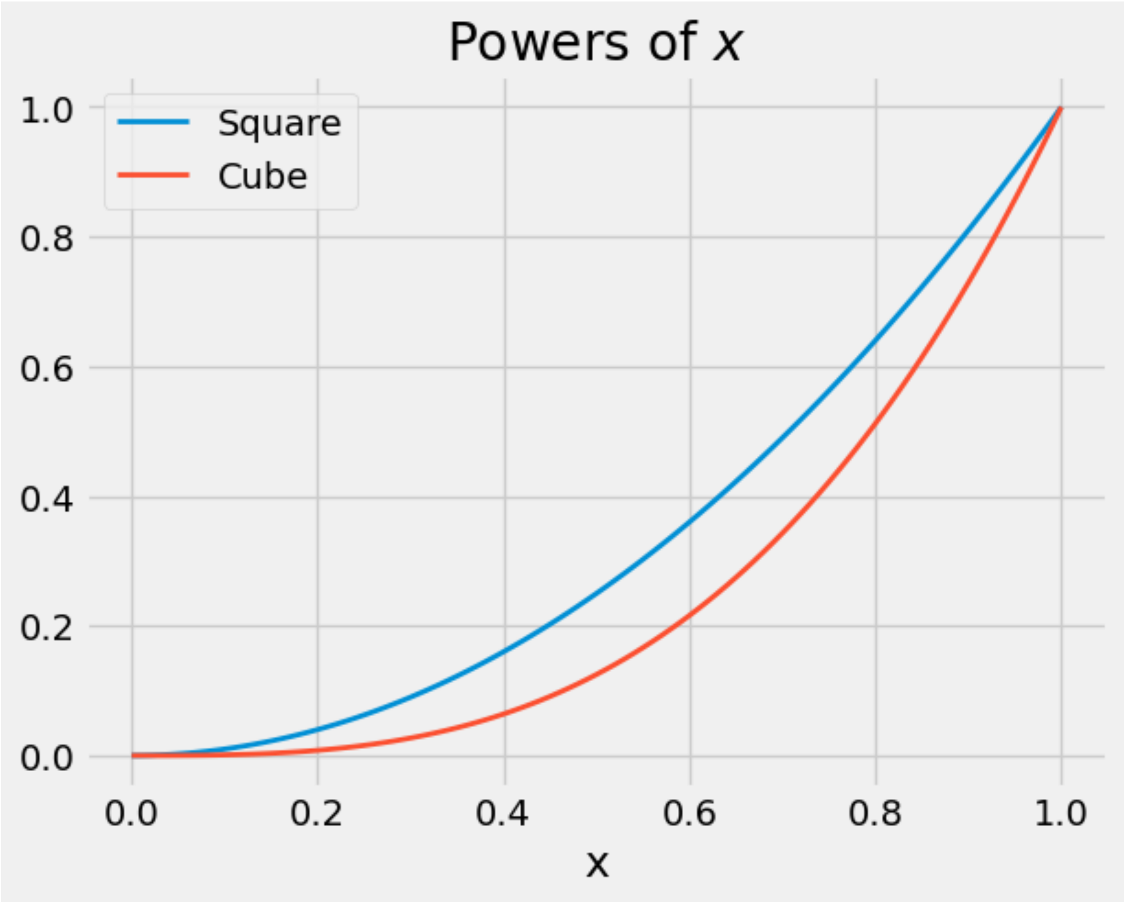
Run these cells to see some examples. Note the overlaid plots: they are straightforward to draw, and Python chooses a different color for each plot.

[2]:
```
x = np.arange(0, 1.01, 0.01)
x_squared = x ** 2
x_cubed = x**3
plt.plot(x, x_squared, lw=2)
plt.plot(x, x_cubed, lw=2);
```



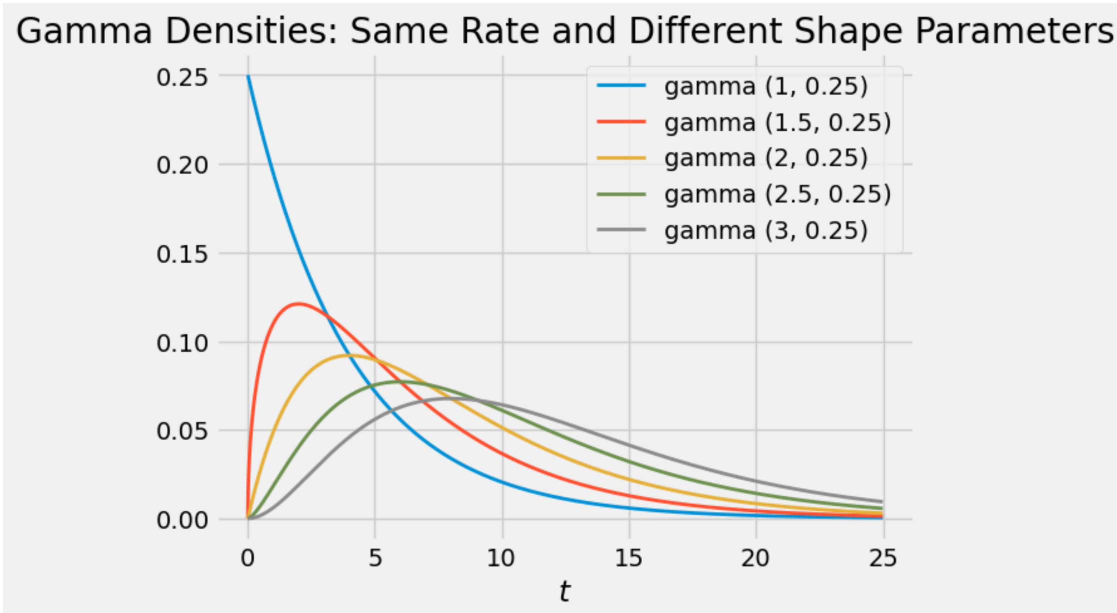
You can add titles, labels, and legends.

[3]:
```
plt.plot(x, x_squared, lw=2, label = 'Square')
plt.plot(x, x_cubed, lw=2, label = 'Cube')
plt.legend()
plt.xlabel('x')
plt.title('Powers of $x$');
```



As another example, the code below was used in Homework 9 Exercise 3d to plot gamma densities.

[4]:
```
t = np.arange(0, 25, 0.01)
r_1 = stats.gamma.pdf(t, 1, scale=1/0.25)
r_1_5 = stats.gamma.pdf(t, 1.5, scale=1/0.25)
r_2 = stats.gamma.pdf(t, 2, scale=1/0.25)
r_2_5 = stats.gamma.pdf(t, 2.5, scale=1/0.25)
r_3 = stats.gamma.pdf(t, 3, scale=1/0.25)
plt.plot(t, r_1, lw=2, label='gamma (1, 0.25)')
plt.plot(t, r_1_5, lw=2, label='gamma (1.5, 0.25)')
plt.plot(t, r_2, lw=2, label='gamma (2, 0.25)')
plt.plot(t, r_2_5, lw=2, label='gamma (2.5, 0.25)')
plt.plot(t, r_3, lw=2, label='gamma (3, 0.25)')
plt.xlabel('$t$')
plt.legend()
plt.title('Gamma Densities: Same Rate and Different Shape Parameters');
```


Gamma Densities: Same Rate and Different Shape Parameters
gamma (1, 0.25)
gamma (1.5, 0.25)
gamma (2, 0.25)
gamma (2.5, 0.25)
gamma (3, 0.25)
0.25
0.20
0.15
0.10
0.05
0.00
0
5
10
15
20
25
t

1.2 1. MLE of the Exponential Rate

For $n > 1$, let $X_1, X_2, \dots, X_n$ be i.i.d. exponential (λ) variables.

a) Let $\hat{\lambda}_n$ be the maximum likelihood estimate (MLE) of the parameter λ. Find $\hat{\lambda}_n$ in terms of the sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. The subscript n in $\bar{X}_n$ is there to remind us that we have the average of n values. It doesn't refer to the nth sampled value X_n.

b) Use facts about sums and linear transformations to find the distribution of $\bar{X}_n$ with little or no calculation. Recognize it as one of the famous ones and provide its name and parameters. Use it to find $E(\hat{\lambda}_n)$.

c) Is $\hat{\lambda}_n$ an unbiased estimate of λ? If it is biased, does it overestimate on average, or does it underestimate? Is it asymptotically unbiased? That is, does $E(\hat{\lambda}_n)$ converge to λ as $n \to \infty$?

1.3 2. Discrete and Continuous Random Selections

Fix a positive integer n, and let p be strictly between 0 and 1.

Suppose Person A picks a number uniformly in the interval $(0, n)$. Let X be the number Person A picks.

Suppose that independently of Person A, Person B picks an integer Y according to the binomial (n, p) distribution, for example by using `stats.binom.rvs(n, p, size=1)` or by tossing a p-coin n times and recording the number of heads.

Find $P(X < Y)$.

1.4 3. MLE and MAP Estimates

The coin is tossed 10 times and the resulting sequence is HTTHHHTHTH. In the parts below, we refer to this information as “the data”.

a) Under the assumption that the coin lands heads with a fixed unknown probability p, find the MLE of p based on the data.

b) Suppose now that the coin lands heads with a random probability X. Let the prior density of X be uniform on the unit interval. Find the MAP estimate of the probability of heads, given the data.

c) Show that if $r > 1$ and $s > 1$ then the mode of the beta (r, s) distribution is $(r-1)/(r+s-2)$. Remember to ignore multiplicative constants and take the log before maximizing.

d) Suppose instead that the prior density of X is $f(x) = 4x^3$ if $0 < x < 1$ and 0 otherwise. Find the MAP estimate of the probability of heads, given the data.

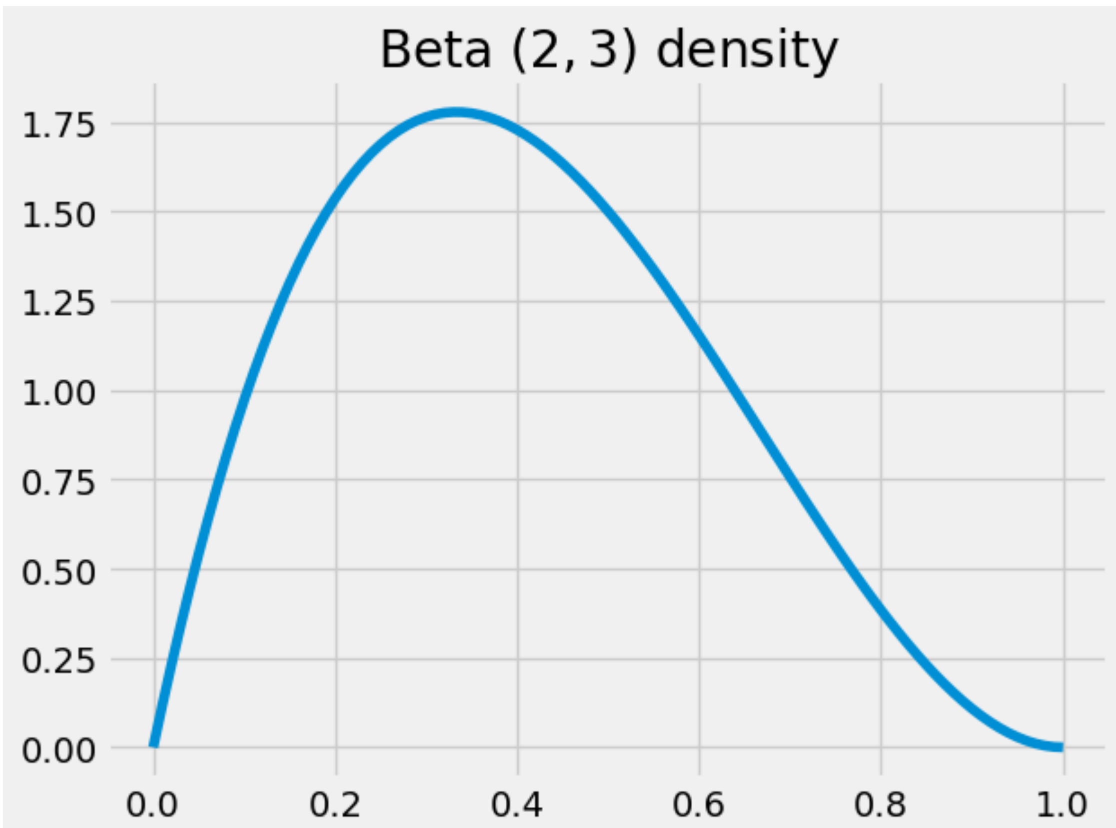
1.5 4. Heads in Tosses of a Random Coin

Let X be a random proportion with a prior distribution that is beta $(2, 3)$. Given $X = p$, let $I_1, I_2, I_3, \ldots$ be i.i.d. Bernoulli (p).

a) Plot the prior density of X.

[5]:
```
# Answer to a

x = np.arange(0, 1.01, 0.01)
y = stats.beta.pdf(x,2,3)
plt.plot(x,y)
plt.title('Beta $(2, 3)$ density');
```



b) Let $H_7 = \sum_{k=1}^{7} I_k$. For each $h = 0, 1, \ldots, 7$, plot the posterior density of X given $H_7 = h$. All 8 plots should be on the same graph.

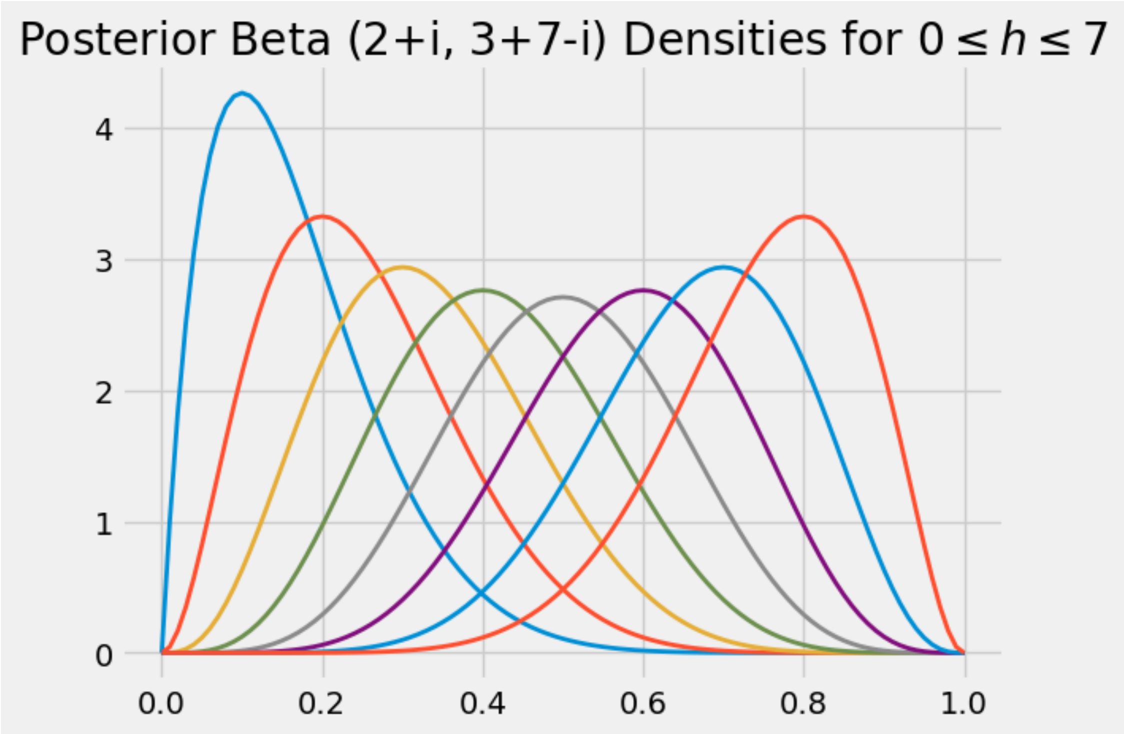
Use as many lines of code as you need. You don't have to include labels and a legend, but the title should say which densities you are plotting.

[6]:
```
# Answer to b

x = np.arange(0, 1.01, 0.01)
```

```python
for i in range(8):
    a = 2 + i
    b = 3 + (7 - i)
    y = stats.beta.pdf(x,a,b)
    plt.plot(x,y,lw=2)

plt.title('Posterior Beta (2+i, 3+7-i) Densities for $0 \leq h \leq 7$');
```



c) What is the MAP estimate of the random probability of heads given $H_7 = 2$? Calculate the estimate using the appropriate formula and confirm that your answer agrees with the estimate visible in the appropriate graph above.

d) Find $P(I_8 = 1 \mid H_7 = 2)$. Your answer should be a decimal value.

e) Find $P(I_8 = 1, I_9 = 1, I_{10} = 1 \mid H_7 = 2)$. Your answer should be a decimal value. Is it equal to the cube of the answer to Part (d)? If not, which is bigger?

1.6 5. Waiting for a Random Coin to Land Heads

Let X be a random proportion. Given $X = p$, let T be the number of tosses till a p-coin lands heads.

a) Let $P(X = 1/10) = 1/4$, $P(X = 1/7) = 1/4$, and $P(X = 1/3) = 1/2$. Find $E(T)$.

b) Find $E(T)$ if X has the beta (r, s) density for some $r > 1$. Simplify all integrals and Gamma functions in your answer.

c) Let X have the beta (r, s) density. For fixed $k > 0$, find the posterior density of X given $T = k$. Identify it as one of the famous ones and state its name and parameters.

1.7 Submission Instructions

Many assignments throughout the course will have a written portion and a code portion. Please follow the directions below to properly submit both portions.

1.7.1 Written Portion

- Scan all the pages into a PDF. You can use any scanner or a phone using applications such as CamScanner. Please **DO NOT** simply take pictures using your phone.
- Please start a new page for each question. If you have already written multiple questions on the same page, you can crop the image in CamScanner or fold your page over (the old-fashioned way). This helps expedite grading.
- It is your responsibility to check that all the work on all the scanned pages is legible.
- If you used $\LaTeX$ to do the written portions, you do not need to do any scanning; you can just download the whole notebook as a PDF via LaTeX.

1.7.2 Code Portion

- Save your notebook using `File > Save and Checkpoint`.
- Generate a PDF file using `File > Download As > PDF via LaTeX`. This might take a few seconds and will automatically download a PDF version of this notebook.
  - If you have issues, please post a follow-up on the general Homework 11 Ed thread.

1.7.3 Submitting

- Combine the PDFs from the written and code portions into one PDF. Here is a useful tool for doing so.
- Submit the assignment to Homework 11 on Gradescope.
- **Make sure to assign each page of your pdf to the correct question.**
- **It is your responsibility to verify that all of your work shows up in your final PDF submission.**

If you are having difficulties scanning, uploading, or submitting your work, please read the Ed Thread on this topic and post a follow-up on the general Homework 11 Ed thread.

[ ]:
```

```

1a. $X_1, X_2, \dots, X_n \sim$ iid Exponential(λ)

$$f(x \mid \lambda) = \lambda e^{-\lambda x} \qquad x \geq 0$$

$$\begin{aligned}
\mathcal{L}ik(\lambda) &= p(X_1, \dots, X_n \mid \lambda) \\
&= p(X_1 \mid \lambda) \times \dots \times p(X_n \mid \lambda) \\
&= \lambda e^{-\lambda x_1} \times \dots \times \lambda e^{-\lambda x_n} \\
\ell(\lambda) &= \log(\lambda e^{-\lambda x_1} \times \dots \times \lambda e^{-\lambda x_n}) \\
&= n \log(\lambda) - \lambda \sum_{i=1}^{n} x_i
\end{aligned}$$

MLE(λ) → need max $\mathcal{L}ik(\lambda)$

∴ derivative $= 0$

$$\frac{d\ell(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

$$0 = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i = \frac{n}{\lambda}$$

$$\lambda = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$= \frac{n \cdot \frac{1}{n}}{\frac{1}{n}\sum_{i=1}^{n} x_i}$$

$$\hat{\lambda}_{mle} = \frac{1}{\bar{x}_n}$$

16. $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

$X_i \sim \text{Exponential}(\lambda)$
$\sim \text{Gamma}(1, \lambda)$

$\sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \lambda)$ textbook 18.3.4

$\bar{X}_n \sim \text{Gamma}(n, \frac{\lambda}{1/n})$
$\sim \text{Gamma}(n, \lambda n)$

$$E[\hat{\lambda}_n] = E\left[\frac{1}{\bar{X}_n}\right]$$
$$= \int_0^\infty \frac{1}{x} f_x(x)\, dx$$
$$= \int_0^\infty \frac{1}{x} \cdot \frac{(\lambda n)^n}{\Gamma(n)} x^{n-1} e^{-\lambda n x}\, dx$$
$$= \int_0^\infty \frac{(\lambda n)^n}{(n-1)!} x^{n-2} e^{-\lambda n x} \cdot \frac{(n-1)}{(n-1)}\, dx$$
$$= \frac{\lambda n}{n-1} \underbrace{\int_0^\infty \frac{(\lambda n)^{n-1}}{(n-2)!} x^{n-2} e^{-\lambda n x}\, dx}_{= \text{gamma}(n-1, \lambda n)}$$
$$= \frac{\lambda n}{n-1} \cdot 1$$
$$= \frac{\lambda n}{n-1}$$

1c. $\hat{\lambda}_n = \frac{1}{\bar{x}_n}$

$E[\hat{\lambda}_n] = \lambda$?

$\frac{\lambda n}{n-1} = \lambda$

$= \frac{\lambda n}{n-1} - \lambda$

$= \lambda \left(\frac{n}{n-1} - 1 \right)$

$= \lambda \left(\frac{n - n + 1}{n-1} \right)$

$= \lambda \underbrace{\left(\frac{1}{n-1} \right)}$, $n > 1$

$\frac{1}{n-1} > 0$

$\therefore E[\hat{\lambda}_n] > \lambda$

$\therefore$ $\hat{\lambda}_n$ is biased and overestimates

$\lim_{n \to \infty} E[\hat{\lambda}_n] = \lim_{n \to \infty} \frac{\lambda n}{n-1}$

$= \lim_{n \to \infty} \frac{\lambda n \cdot \frac{1}{n}}{\frac{1}{n}(n-1)}$

$= \lim_{n \to \infty} \frac{\lambda}{1 - \cancel{\frac{1}{n}}_{0}}$

$= \lambda$

$\therefore$ $\hat{\lambda}_n$ is asymptotically unbiased

2. $X \sim \text{Uniform}(0, n)$

$$P(X \le k) = \frac{k}{n}$$

independent

$Y \sim \text{Binomial}(n, p)$

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(X < Y) = \sum_{k=0}^{n} \underbrace{P(X < Y \mid Y = k)}_{= P(X < k)} P(Y = k)$$

$$= \sum_{k=0}^{n} P(X < k) P(Y = k)$$

$$= \sum_{k=0}^{n} \frac{k}{n} \cdot P(Y = k)$$

$$= \sum_{k=1}^{n} \frac{k}{n} \cdot P(Y = k)$$

$$= \frac{1}{n} \sum_{k=1}^{n} \underbrace{k P(Y = k)}_{= E[Y]}$$

$$= \frac{1}{n} \cdot np$$

$$= p$$

3a. HTTHHHTHTH

$H = 6$
$T = 4$

$$\mathcal{L}ik(p) = p^6(1-p)^4$$
$$\ell(p) = \log(p^6(1-p)^4)$$
$$= 6\log(p) + 4\log(1-p)$$
$$\frac{d\,\ell(p)}{dp} = \frac{6}{p} + \frac{4}{1-p} \cdot -1 = 0$$
$$\frac{4}{1-p} = \frac{6}{p}$$
$$4p = 6 - 6p$$
$$10p = 6$$
$$p = \frac{3}{5}$$

3b. X : prior : uniform $(0,1)$ = beta $(1,1)$
posterior : beta $(6+1, 4+1)$
beta $(7,5)$

$$\frac{r-1}{r+s-2} = \frac{7-1}{7+5-2}$$ textbook 20.3.3
$$= \frac{6}{10}$$
$$= \frac{3}{5}$$

3c. $f(p) \propto p^{r-1}(1-p)^{s-1}$
$$\log(f(p)) = \log(p^{r-1}(1-p)^{s-1})$$
$$= (r-1)\log(p) + (s-1)\log(1-p)$$
$$\frac{d\,f(p)}{dp} = \frac{r-1}{p} + \frac{s-1}{1-p} \cdot -1 = 0$$
$$\frac{s-1}{1-p} = \frac{r-1}{p}$$

$$p(s-1) = (1-p)(r-1)$$

$$ps - p = r - 1 - pr + p$$

$$pr + ps - 2p = r - 1$$

$$p(r+s-2) = r-1$$

$$p = \frac{r-1}{r+s-2}$$

3d. X : prior density : $f(x) = 4x^3$

$f(p) = 4p^3$

same likelihood : $\mathcal{L}ik(p) = p^6(1-p)^4$

posterior $\propto$ prior $\times$ likelihood textbook 20.3.1

$$f(p \mid \text{data}) \propto 4p^3 \cdot p^6 (1-p)^4$$

$$= 4p^9 (1-p)^4$$

$$\therefore \text{Beta}(10, 5)$$

$$\frac{r-1}{r+s-2} = \frac{10-1}{10+5-2}$$

$$= \frac{9}{13}$$

4c. X : prior : Beta$(2, 3)$

$H_7 = 2$

$\therefore$ 7 tosses $\rightarrow$ 2 Heads
5 Tails

posterior : Beta$(2+2, 3+5)$
Beta$(4, 8)$

$$\frac{r-1}{r+s-2} = \frac{4-1}{4+8-2}$$

$$= \frac{3}{10}$$

4d. $P(\underbrace{I_8 = 1}_{\text{8th toss lands heads}} \mid H_7 = 2) = E[X \mid H_7 = 2]$

$$E[X] = \frac{r}{r+s}$$ textbook 17.4.5

$$\therefore E[X \mid H_7 = 2] = \frac{4}{4+8}$$

$$= \frac{4}{12}$$

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

4e. $P(\underbrace{I_8 = 1, I_9 = 1, I_{10} = 1}_{\text{tosses 8, 9, 10 ALL land heads}} \mid H_7 = 2) = E[X^3 \mid H_7 = 2]$

$$f_{X \mid H_2 = 7}(x) = \frac{1}{\frac{\Gamma(4)\Gamma(8)}{\Gamma(12)}} \cdot x^3 (1-x)^7$$

$$= \frac{11!}{7!\,3!} x^3 (1-x)^7$$

$$= 1320\, x^3 (1-x)^7$$

$$\therefore E[X^3 \mid H_7 = 2] = \int_0^1 x^3 \cdot 1320\, x^3 (1-x)^7 \, dx$$

$$= 1320 \underbrace{\int_0^1 x^6 (1-x)^7 \, dx}_{= \text{Beta}(7,8)}$$

$$= 1320 \times \frac{\Gamma(7)\Gamma(8)}{\Gamma(15)}$$

$$= 1320 \times \frac{6!\,7!}{14!}$$

$$\approx 0.0549$$

$$(E[X \mid H_7 = 2])^3 = \left(\frac{1}{3}\right)^3$$

$$\approx 0.037$$

$$\therefore E[X^3 \mid H_7 = 2] \neq (E[X \mid H_7 = 2])^3$$

$$0.0549 > 0.037$$

5a. $T \sim$ # tosses till a p-coin lands heads
$\sim$ Geometric(p)

$$p(X = \tfrac{1}{10}) = \tfrac{1}{4}$$
$$p(X = \tfrac{1}{7}) = \tfrac{1}{4}$$
$$p(X = \tfrac{1}{3}) = \tfrac{1}{2}$$

$$E[T \mid X = \tfrac{1}{10}] = \frac{1}{\frac{1}{10}} = 10$$

$$E[T \mid X = \tfrac{1}{7}] = \frac{1}{\frac{1}{7}} = 7$$

$$E[T \mid X = \tfrac{1}{3}] = \frac{1}{\frac{1}{3}} = 3$$

$$E[T] = \tfrac{1}{4}(10) + \tfrac{1}{4}(7) + \tfrac{1}{2}(3) = \frac{10 + 7 + 6}{4} = \frac{23}{4}$$

5b. $X \sim$ Beta(r, s)

$$E[T] = E[\tfrac{1}{X}] = \int_0^1 \frac{1}{p} f_X(p)\, dp = \int_0^1 \frac{1}{p} \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} p^{r-1} (1-p)^{s-1}\, dp$$

$$= \int_0^1 \frac{1}{p} \cdot \frac{(r+s-1)!}{(r-1)!(s-1)!} p^{r-1}(1-p)^{s-1}\, dp$$

$$= \frac{(r+s-1)!}{(r-1)!\,(s-1)!} \underbrace{\int_0^1 p^{r-2}(1-p)^{s-1}\, dp}_{:\ \text{Beta}(r-1,\ s)}$$

$$= \frac{(r+s-1)!}{(r-1)!\,(s-1)!} \cdot \frac{\Gamma(r-1)\Gamma(s)}{\Gamma(r-1+s)}$$

$$= \frac{(r+s-1)!}{(r-1)!(s-1)!} \cdot \frac{(r-2)!(s-1)!}{(r+s-2)!}$$

$$= \frac{r+s-1}{r-1}$$

5c. $T \sim$ # tosses till first head

$\sim$ Geometric(p)

$X \sim \text{Beta}(r, s)$

likelihood: $P(T = k \mid X = p) = p\,(1-p)^{k-1}$

prior: $f_X(p) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} p^{r-1}(1-p)^{s-1}$

$$f_{X|T=k}(p) \propto p(1-p)^{k-1} \cdot p^{r-1}(1-p)^{s-1}$$
$$= p\,(1-p)^{s+k-2}$$

$\therefore$ Beta$(r+1,\ s+k-1)$