Homework 6

YOUR NAME   March 28, 2018

- **Acknowledgments:** This template takes some materials from course CSE 547/Stat 548 of Washington University: https://courses.cs.washington.edu/courses/cse547/17sp/index.html.

  If you refer to other materials in your homework, please list here.

- **Collaborators:** I finish this template by myself. If you finish your homework all by yourself, make a similar statement. If you get help from others in finishing your homework, state like this:

  - 1.2 (b) was solved with the help from _____.
  - Discussion with _____ helped me finishing 1.3.

You may use `enumerate` to generate answers for each question:

6.1. Type of commonly used notations. Use another `enumerate` to start generate answers for sub-questions:

  (a) Use `$ $` to get an inline equation: $\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A(\omega)]$.

  (b) Use `equation` to have equation in display math mode:

  $$\frac{a+b}{2} \geq \sqrt{ab} \tag{1}$$

  (c) Use `\eqref` to get reference for equations: (1) holds when $a \geq 0, b \geq 0$.

  (d) Now we would introduce some commonly used notations:

   i. Use `\mathbb{P}, \mathbb{R}, \mathbb{E}` to type $\mathbb{P}, \mathbb{R}, \mathbb{E}$.
   ii. Use `\mathcal{A}, \mathcal{X}, \mathcal{Y}, \mathcal{N}` to type $\mathcal{A}, \mathcal{X}, \mathcal{Y}, \mathcal{N}$.
   iii. Use `\underline{x}, \underline{y}` to type vectors $\underline{x}, \underline{y}$.
   iv. Use `\mathsf{x}, \mathsf{y}, \mathsf{z}` to type random variables $\mathsf{x}, \mathsf{y}, \mathsf{z}$. For simplicity, I have defined several macros so you could simply type `\rvx, \rvy, \rvz`. Don't forget `$ $`!
   v. Thanks to these macros, we could have $\mathbb{R}, \mathbb{E}[\mathsf{x}], \mathrm{Var}(\mathsf{y}), \mathbb{P}(A), \perp\!\!\!\perp, \mathbb{1}$ by typing `\reals, \E[\rvx], \Var(\rvy), \Prob(A), \independent, \1`.
   vi. Now you can use `\ux, \uy, \uz` to type vectors $\underline{x}, \underline{y}, \underline{z}$, and use `\urvx, \urvy, \urvz` to type random vectors $\underline{\mathsf{x}}, \underline{\mathsf{y}}, \underline{\mathsf{z}}$.

vii. Remember that $P_{x|y}(x|y) \triangleq \mathbb{P}(x = x|y = y)$.

  α) Writing $\mathbb{P}(x)$ is wrong. $\mathbb{P}$ should only operate on events.

  β) x is a random variable, while x is a real number.

(e) You may find `https://en.wikibooks.org/wiki/LaTeX` useful.

(f) Writing LaTeX online may be easier for beginners:

  i. ShareLaTeX: `https://www.sharelatex.com/`.

  ii. Overleaf: `https://www.overleaf.com/`.

6.2. You may need aligned equations for your homework, here are several examples:

Total propability rule:

$$\begin{aligned}
\mathbb{P}(x = x) &= \sum_{y \in \mathcal{Y}} \mathbb{P}(x = x, y = y) \\
&= \sum_{y \in \mathcal{Y}} \mathbb{P}(x = x|y = y) \, \mathbb{P}(y = y),
\end{aligned}$$

or

$$\begin{aligned}
P_x&(x) \\
&= \sum_{y \in \mathcal{Y}} P_{xy}(x, y) \\
&= \sum_{y \in \mathcal{Y}} P_{x|y}(x|y) P_y(y).
\end{aligned}$$

Indicator function:

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A, \\ 0, & \text{if } \omega \notin A. \end{cases}$$

6.3. You may need to add figure and source codes in your homework. Figure 1 is an example that compares the empirical distribution (histogram) and probability density function of the Gaussian random variable.

The source code to plot Figure 1 could be found in Appendix A. Here are the core codes:

```
4  [cnt, x_hist] = hist(data, nbins); % not to plot, only to get
       emperical distribution.
```

```
6  cnt = cnt / n / (x_hist(2) − x_hist(1)); % normalization, be
       careful :)
7  bar(x_hist, cnt); % plot the hist using bar()
```

To understand line 6, note that if we have n samples of x denoted by $x^{(i)}, i = 1, 2, \cdots, n$, then the probability density function p_x could be
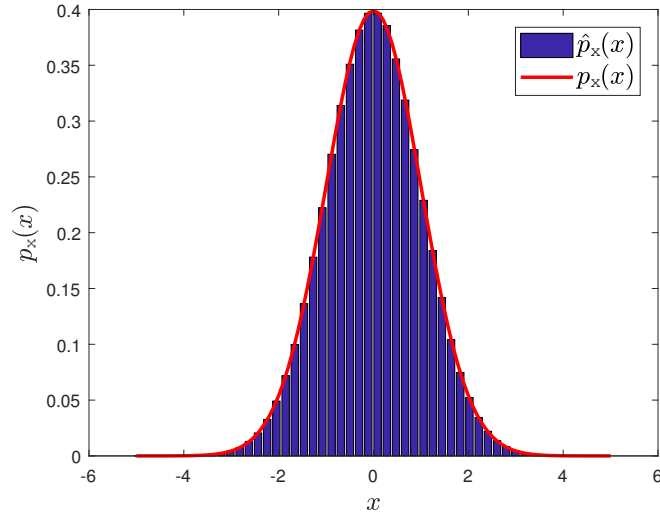
Figure 1: Gaussian PDF and histogram of samples

estimated as

$$
\begin{aligned}
p_{\mathsf{x}}(x_0) &= \left.\frac{\mathrm{d}}{\mathrm{d}x}\,\mathbb{P}(\mathsf{x} \le x)\right|_{x=x_0} \\
&\approx \frac{\mathbb{P}(x_0 - \Delta x < \mathsf{x} \le x_0)}{\Delta x} \\
&\approx \frac{1}{n\Delta x}\sum_{i=1}^{n}\mathbb{1}_{x^{(i)}\in(x_0-\Delta x,\,x_0]}\,.
\end{aligned}
$$

6.4. An example of hypothesis testing:

$$
\log\frac{\mathbb{P}(\mathsf{H}=H_1|\mathsf{y}=y)}{\mathbb{P}(\mathsf{H}=H_0|\mathsf{y}=y)}\mathop{\substack{\hat{\mathsf{H}}=H_1 \\ \gtrless \\ \hat{\mathsf{H}}=H_0}}\gamma
$$

A   Source code

Source code for plotting Figure 1 is shown as follows.

```matlab
n = 1e6;  % n samples
data = randn(1e6, 1); % Generate n Random Gaussian samples.
nbins = 50; % bins in your histogram
[cnt, x_hist] = hist(data, nbins); % not to plot, only to get
    emperical distribution.
figure;
cnt = cnt / n / (x_hist(2) - x_hist(1)); % normalization, be
    careful :)
bar(x_hist, cnt); % plot the hist using bar()
hold on;
x = -5 : 0.1 : 5;
plot(x, normpdf(x), 'r', 'linewidth', 2);
legend({'$\hat{p}_{\sf{x}}(x)$', '$p_{\sf{x}}(x)$'}, 'Interpreter',
     'LaTeX', 'fontsize', 15);
xlabel('$x$', 'Interpreter', 'LaTeX', 'fontsize', 15); % You may
    change the size accordingly
ylabel('$p_{\sf{x}}(x)$', 'Interpreter', 'LaTeX', 'fontsize', 15);
title(⟨your-title-here⟩)
```

Listing 1: FigurePlot