Homework 8 (due Thursday, June 20th at 11:59pm)

1. (10 pts) Let $X_1, \dots, X_n \sim F_X$ where F_X is the $X \sim \text{Beta}(\alpha, \beta)$ with $\alpha = 1, \beta = 2$. Let $n = 200$. Find an approximation for

$$P\left(\log(\bar{X}_n) \leq -\frac{1}{2}\right).$$

Solution:

$$P\left(\bar{X}_n \leq e^{-\frac{1}{2}}\right) = P\left(\sqrt{200} \cdot 3\sqrt{2} \cdot \left(\bar{X}_n - \frac{1}{3}\right) \leq \sqrt{200} \cdot 3\sqrt{2}\left(e^{-\frac{1}{2}} - \frac{1}{3}\right)\right)$$
$$\approx \phi\left(3\sqrt{400}\left(e^{-\frac{1}{2}} - \frac{1}{3}\right)\right) \approx \phi(16.39184) \approx 1.$$

2. (10 pts) Let $V_1, \dots, V_n \overset{iid}{\sim} \text{Beta}(1, \theta)$. Let $W_1 = V_1,\ W_2 = V_2(1 - V_1), \dots,\ W_n = V_n \prod_{i=1}^{n-1}(1 - V_i)$.

- Prove that $1 - V_i \sim \text{Beta}(\theta, 1)$.
- Prove that $E[1 - \sum_{i=1}^{n} W_i] \to 0$ as $n \to \infty$.[1]
  Hint: To prove this, rewrite the $1 - \sum_{i=1}^{n} W_i$ in terms of the V_i's. In order to do this, notice the following pattern:

$$1 - W_1 = 1 - V_1$$
$$1 - W_1 - W_2 = 1 - V_1 - V_2(1 - V_1) = (1 - V_1)(1 - V_2)$$
$$\dots$$

Solution:

- Using the method of change of variables we obtain

$$f_Y(y) = \theta y^{\theta - 1}$$

  from which we can conclude that $Y \sim \text{Beta}(\theta, 1)$.
- By induction you can show that

$$1 - \sum_{i=1}^{n} W_i = \prod_{i=1}^{n}(1 - V_i).$$

[1] This is called *convergence in mean.*

Then

$$E\left[\prod_{i=1}^{n}(1-V_i)\right] = \prod_{i=1}^{n} E[1-V_i] = \left(\frac{\theta}{1+\theta}\right)^n \overset{n\to\infty}{\longrightarrow} 0.$$

This is the first (informal) part of the stick-breaking construction of the Dirichlet process.

3. (20 pts) Let $X_1, \dots, X_n \overset{iid}{\sim} F_X$ where

$$F_{(X_1,\dots,X_n)}(x_1, \dots, x_n) = \prod_{i=1}^{n}(1-e^{-x_i})\mathbb{1}(x_i > 0).$$

- Derive $f_{X_1}(x)$.
- Show that the m.g.f. for X_1 is

$$m_{X_1}(t) = \frac{1}{1-t}$$

  for $t \in (-\epsilon, \epsilon)$. What is the value of ϵ?
- Show that the m.g.f. of $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ is

$$m_{\bar{X}_n}(t) = \left[m_{X_1}\left(\frac{t}{n}\right)\right]^n$$

  again, $\forall t/n \in (-\epsilon, \epsilon)$ as in the example above.
- Show that

$$P\left(\sum_{i=1}^{n}\frac{X_i}{n} > a\right) \le e^{-2t}$$

  for $a = 3$.
  Hint: You do not really need the previous step for this proof, but you can use it.

Solution:

- First, notice that

$$F_{(X_1,\dots,X_n)}(x_1, \dots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i) \text{ by independence}$$

therefore we can now derive f_{X_1} from F_{X_1}.

$$f_{X_1}(x) = \frac{\partial F_{X_1}(x)}{\partial x} = e^{-x}\mathbb{1}(x > 0).$$

This means that $X_1 \sim \text{Exponential}(1)$.

- $$m_X(t) = E[e^{tX}] = \int_0^\infty e^{tx}e^{-x}dx = \int_0^\infty e^{x(t-1)}dx$$
$$= \left.\frac{e^{x(t-1)}}{t-1}\right|_0^\infty = \begin{cases} \frac{1}{1-t} \text{ if } t < 1 \\ \infty \text{ o/w} \end{cases}$$

  therefore we conclude that $\epsilon = 1$, that is $t \in (-1, 1)$.

- $$m_{\bar{X}_n}(t) = E\left[e^{t\bar{X}_n}\right] = E\left[e^{\frac{t}{n}\sum_{i=1}^n}\right]$$
$$= \prod_{i=1}^n E\left[e^{\frac{t}{n}X_i}\right] = \left(E\left[e^{\frac{t}{n}X_1}\right]\right)^n$$
$$= \left[m_{X_1}\left(\frac{t}{n}\right)\right]^n = \left[\frac{1}{1-\frac{t}{n}}\right]^n = \left(\frac{n}{n-t}\right)^n.$$

  where the second to last inequality is thanks to the fact that the r.v.'s are identically distributed. This holds for $t/n \in (-1, 1)$.

- For $a \in \mathbb{R}$ and $t \geq 0$,

$$P(\bar{X}_n \geq a) = P(t\bar{X}_n \geq ta)$$
$$= P\left(e^{t\bar{X}_n} \geq e^{ta}\right) \leq \frac{E\left[e^{t\bar{X}_n}\right]}{e^{ta}} \text{ by Markov's}$$
$$= \left(\frac{n}{n-t}\right)^n \frac{1}{e^{ta}} = \left(1 - \frac{t}{n}\right)^{-n} e^{-ta} \overset{n\to\infty}{\longrightarrow} \frac{e^t}{e^{ta}}$$

  Hence for $a = 3$, we obtain, for $n \to \infty$,

$$P(\bar{X}_n \geq 3) \leq e^{-2t}$$

4. (15 pts) Assume that $X_1, \cdots, X_n$ have *finite* moment generating functions; i.e. $m_{X_i}(t) < \infty$ for all t, for all $i = 1, \cdots, n$. Prove the central

limit theorem using moment generating functions.
Hint: consier the random variable

$$T_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Then study the mgf of T_n, and its convergence: you need to prove that it converges to $e^{\frac{t^2}{2}}$, since this is the mgf of a standard normal. Remember that for a standard normal all even central moments are zero!

Solutions:
Let $T_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, for the sum $S_n = \sum_i X_i$. Also denote the mean and variance of X as $E(X_i) = \mu$ and $V(X_i)\sigma^2$. Then, we want to prove that

$$P(T_n \le x) \overset{n\to\infty}{\to} P(Z \le x)$$

Denote as the 'standardized' version of X_i as $Y_i = \frac{X_i - \mu}{\sigma}$ – we see that Y_i are still i.i.d. but with mean $E(Y_i) = 0$ and $V(Y_i) = 1$, and

$$T_n = \frac{\sum_{i=1}^n Y_i}{\sqrt{n}}$$

Now, the central limit theorem is to show that the moment generating function of T_n goes to that of a standard normal random variable $Z \sim \mathcal{N}(0,1)$, whose mgf is $\exp(\frac{t^2}{2})$.

$$\begin{aligned}
m_{T_n}(t) &= E(e^{t \cdot T_n}) \\
&= \left(E(e^{\frac{t}{\sqrt{n}} Y_i})\right)^n \\
&= \left(1 + \frac{t}{\sqrt{n}} EY + \frac{1}{2}\frac{t^2}{n} E(Y^2) + \frac{1}{6}\frac{t^3}{n^{3/2}} E(Y^3) + \cdots\right)^n \\
&= \left(1 + 0 + \frac{1}{2}\frac{t^2}{n} + \frac{1}{6}\frac{t^3}{n^{3/2}} E(Y^3) + \cdots\right)^n \\
&\approx \left(1 + \frac{t^2}{2}\frac{1}{n}\right)^2 \\
&\to e^{\frac{t^2}{2}}
\end{aligned}$$

5. (15 pts) Let $X_i \sim \text{Gamma}(\alpha_i, \beta)$[2] for $i = 1, \ldots, n$, and $X_i \perp\!\!\!\perp X_j$ for $1 \le i < j \le n$.

[2] β is the rate parameter.

- Prove that the mgf of X_1 is

$$m_{X_1}(t) = \left(1 - \frac{t}{\beta}\right)^{-\alpha}$$

  for $t < \beta$.
- Prove that $S_n = X_1 + \cdots + X_n \sim \text{Gamma}(\sum_{i=1}^n \alpha_i, \beta)$.
- Conclude that $S_n = Z_1^2 + \cdots + Z_n^2 \sim \chi^2(n)$ where $Z_1, \ldots, Z_n \overset{iid}{\sim} \mathcal{N}(0,1)$.
  Hint: you can derive it directly or, better, remember the relationship of this distribution with the Gamma....

Solution:

- $$\begin{aligned} m_{X_1}(t) = E[e^{tX}] &= \int_0^\infty \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1 - 1} e^{-(\beta - t)x} dx \\ &= \frac{\beta^{\alpha_1}}{(\beta - t)^{\alpha_1}} \int_0^\infty \frac{(\beta - t)^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1 - 1} e^{-(\beta - t)x} dx \end{aligned}$$

  hence for $t < \beta$ we have

$$= \frac{\beta^{\alpha_1}}{(\beta - t)^{\alpha_1}} = \left(\frac{\beta}{\beta - t}\right)^{\alpha_1} = \left(1 - \frac{t}{\beta}\right)^{-\alpha}.$$

- $$E\left[e^{t \sum_{i=1}^n X_i}\right] = \prod_{i=1}^n \left(1 - \frac{t}{\beta}\right)^{-\alpha_i} = \prod_{i=1}^n \left(1 - \frac{t}{\beta}\right)^{-\alpha_i} = \left(1 - \frac{t}{\beta}\right)^{-\sum_{i=1}^n \alpha_i}.$$

- Remember that a $Z \sim \chi^2(n)$ is equivalent to $Z \sim \text{Gamma}(n/2, 1/2)$. Therefore, plugging these values into the mgf for the Gamma distribution, we obtain

$$(1 - 2t)^{-\frac{n}{2}}.$$

  for $t < 1/2$.

6. (20 pts) Now we'll do something useful for the next time you go to the casino. There are two slot machines: one is giving the prize \$1 with probability p_1, while the second machine is giving the prize \$1

with probability p_2. To make the problem interesting assume that $1 > p_1 > 0.5 > p_2 > 0$. You want to become rich, so you want to choose machine (1)! Of course you do not observe the p's (the true probabilities), so you start playing on both machines the same amount of times.The most naive strategy would be to choose the machine with the highest mean, but ... How likely is it that you'll make the right choice?

Let's work it out. Let $X_i \in \{0,1\}$ be the i-th results of machine 1, and $Y_i \in \{0,1\}$ the i-th result of machine 2. Assume the machines to be independent. Compute the quantities for $p_1 = 0.6, p_2 = 0.4$.

- You play only once. In case of $Y_1 = X_1 = 1$ or $X_1 = Y_1 = 0$ you choose the machine randomly. How likely is it that you'll pick the right machine? *Hint:* compute $P(X_1 > Y_1)$.
- Let's just consider the normal approximation to make everything easier. What is $P(\bar{X}_n > \bar{Y}_n)$? Compute it for $n = 30$, $n = 200$, and for $n = 10^6$. How does this probability compare to $n = 1$? For simplicity do not take into account the case of ties in the latter.
  Hint: remember the property of the normals.
- Now let's see how "fast" this probability converges to 1. Find a lower bound to $P(\bar{X}_n > \bar{Y}_n)$ in terms of n.
  Hint: Use the following facts:
  - $P(X \geq a) \leq b \iff P(X < a) \geq 1 - b$.
  - $P(Z > x) \leq e^{-\frac{x^2}{2}}/(x\sqrt{2\pi})$ for $Z \sim \mathcal{N}(0,1)$.

Solution:

- $$P(X_1 > Y_1) = P(X_1 - Y - 1 > 0) = p_{X,Y}(1,0) = p_X(1)p_Y(0) = p_1(1-p_2)$$

  and
  $$P(X_1 = Y_1) = p_1p_2 + (1-p_1)(1-p_2).$$

  Therefore I will choose the correct machine with probability

  $$p_1(1-p_2) + \frac{p_1p_2 + (1-p_1)(1-p_2)}{2} = 0.6.$$

- Remember that $\bar{X}_n - \bar{Y}_n \xrightarrow{d} \mathcal{N}\left(p_1 - p_2, \frac{1}{n}(p_1(1-p_1) + p_2(1-p_2)\right)$ by CLT and properties of the normal distribution.

$$P(\bar{X}_n > \bar{Y}_n) = P(\bar{X}_n - \bar{Y}_n > 0)$$

$$= P\left(\sqrt{n}\frac{\bar{X}_n - \bar{Y}_n - p_1 + p_2}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}} > \sqrt{n}\frac{p_2 - p_1}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}\right)$$

$$= 1 - \phi\left(\sqrt{n}\frac{p_2 - p_1}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}\right)$$

For $n = 30$ we obtain 0.9430769, for $n = 200$ we have 0.9999777, and for $n = 10^6$ we obtain 1!

- As in the previous exercise, we have

$$P\left(\bar{X}_n > \bar{Y}_n\right) = \phi\left(\sqrt{n}\frac{p_1 - p_2}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}\right)$$

where we use the symmetricity around 0 of the standard normal. Now, let's recall the hint

$$P(X \geq a) \leq b \iff P(X < a) \geq 1 - b.$$

It's easy to notice that with this hint we can lower bound the formula above using the other hint:

$$\phi\left(\sqrt{n}\frac{p_1 - p_2}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}\right) \geq 1 - \frac{1 - \exp\left\{-\frac{1}{2}x^2\right\}}{x\sqrt{2\pi}}$$

where

$$x = \sqrt{n}\frac{p_1 - p_2}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}.$$

Call

$$a := \frac{p_1 - p_2}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}$$

and now we can rewrite the results as

$$\phi(\sqrt{n}a) \geq 1 - \frac{1 - \exp\left\{-\frac{a^2}{2}n\right\}}{\sqrt{n}a\sqrt{2\pi}}.$$

7. (10 pts) Prove/disprove the following statements:

(a) Let $X \sim \mathcal{N}(0,1)$. Is X/c where $c \in \mathbb{R}$ normally distributed? What about cX?

(b) Let $X_1, X_2 \sim \text{Uniform}(0,1)$. Can you compute $P(X_1 + X_2 \geq 1)$? Can you upper/lower bound it?

(c) $X_n \sim \mathcal{N}(0, \sigma_n^2)$ where σ_n^2 is 1 if n is odd, and 2 o/w. Does X_n convergence in distribution to some distribution P?

Solution:

(a) Yes.

(b) You can only lower/upper bound it because they might not be independent.

(c) No.