

MATH 241 Fall 2016 Homework #8

Professor Adam Kapelner

Due 2PM, Friday, December 2, 2016

(this document last updated Thursday 24th November, 2016 at 1:51pm)

Instructions and Philosophy

The path to success in this class is to do many problems. Unlike other courses, exclusively doing reading(s) will not help. Coming to lecture is akin to watching workout videos; thinking about and solving problems on your own is the actual “working out”. Feel free to “work out” with others; **I want you to work on this in groups.**

Reading is still *required*. For this homework set, read moment generating functions and their properties and read about the central limit theorem.

The problems below are color coded: **green** problems are considered *easy* and marked “[easy]”; **yellow** problems are considered *intermediate* and marked “[harder]”, **red** problems are considered *difficult* and marked “[difficult]” and **purple** problems are extra credit. The *easy* problems are intended to be “giveaways” if you went to class. Do as much as you can of the others; I expect you to at least attempt the *difficult* problems.

This homework is worth 100 points but the point distribution will not be determined until after the due date. See syllabus for the policy on late homework.

Up to 15 points are given as a bonus if the homework is typed using L^AT_EX. Links to installing L^AT_EX and program for compiling L^AT_EX is found on the syllabus. You are encouraged to use overleaf.com. If you are handing in homework this way, upload `hwxx.tex` and `preamble.tex`, read the comments in the code; there are two lines to comment out and you should replace my name with yours and write your section. If you are asked to make drawings, you can take a picture of your handwritten drawing and insert them as figures or leave space using the “`\vspace`” command and draw them in after printing or attach them stapled.

The document is available with spaces for you to write your answers. If not using L^AT_EX, print this document and write in your answers. I do not accept homeworks not on this printout. Keep this first page printed for your records. Write your name and section below (A or B).

NAME: _____ SECTION (A or B): _____

Moment Generating Functions There are a few facts to know for the exam. This is what you are responsible for.

Problem 1

In this problem you will be introduced to mgf's and learn the three facts about them that we will make use of (1) that mgf's can give you moments easily and (2) it is easy to obtain the mgf of convolutions of r.v.'s. (3) r.v.'s with equivalent mgf's have equivalent CDF's, making them identically distributed.

- (a) [easy] What is a moment generating function (mgf)? Read about it in the book and answer a few sentences *in English*.

- (b) [easy] Using the definition of expectation, write out what $\mathbb{E}[g(X)]$ is for a discrete r.v. X . This can be found in the notes from previous lectures.

- (c) [easy] Using the definition of expectation, write out what $\mathbb{E}[g(X)]$ is for a continuous r.v. X . This can be found in the notes from previous lectures.

- (d) [easy] For the mgf, what is the $g(X)$ transformation we are interested in finding the expectation of? Write $g(X) =$ something below.

- (e) [easy] If X is a discrete r.v., write the definition of the mgf, $M_X(t)$. If you use the notation $f(x)$ make sure you indicate what you are referring to. If you use the notation $\mathbb{E}[\cdot]$, make sure you indicate the mathematical definition of the expectation.

- (f) [easy] If X is a continuous r.v., write the definition of the mgf, $M_X(t)$. If you use the notation $f(x)$ make sure you indicate what you are referring to. If you use the notation $\mathbb{E}[\cdot]$, make sure you indicate the mathematical definition of the expectation.
- (g) [easy] Using the Taylor series definition of the exponential function, write out the first five terms of what e^{tX} is and then write $+ \dots$ afterwards.
- (h) [easy] Using your answer from part (g), write the first five terms of $\mathbb{E}[e^{tX}]$ using the property of expectation being a linear operator (*i.e.* $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ which is valid always) and then write $+ \mathbb{E}[\dots]$ afterwards.
- (i) [easy] You have written the first five terms of the infinite series representation of $M_X(t)$, the mgf for a r.v. X in the previous problem. Now find $M'_X(t)$, the derivative of the mgf with respect to t , the dummy variable in the t domain and then write $+ \dots$ afterwards. Remember, anything that's a function of the r.v. X is a constant when differentiating with respect to t .
- (j) [easy] Find $M'_X(0)$, *i.e.* the derivative of the mgf evaluated at $t = 0$. Ignore terms after the ...

- (k) [easy] Find $M_X''(t)$, the second derivative of the mgf with respect to t and then write $+\dots$ afterwards.
- (l) [easy] Find $M_X''(0)$, *i.e.* the second derivative of the mgf evaluated at $t = 0$. Ignore terms after the ...
- (m) [easy] Find $M_X'''(t)$, the third derivative of the mgf with respect to t and then write $+\dots$ afterwards.
- (n) [easy] Find $M_X'''(0)$, *i.e.* the third derivative of the mgf evaluated at $t = 0$. Ignore terms after the ...
- (o) [easy] When evaluating parts (j), (l), (n) at $t = 0$, why do you think you're entitled to ignore terms after the ...? Write a sentence *in English*.
- (p) [easy] Synthesize your answers to the last 7 parts and explain why the mgf is called the "moment generating" function *in English*.

- (q) [easy] Let X_1 and X_2 be two independent r.v.'s (not necessarily identically distributed). Show that $M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t)$.

- (r) [harder] Generalize this reasoning and show that for X_1, X_2, \dots, X_n , a series of independent r.v.'s discrete or continuous and $T_n = \sum_{i=1}^n X_i$ that the following is true:

$$M_{T_n}(t) = \prod_{i=1}^n M_{X_i}(t)$$

- (s) [difficult] In (q) and (r), why do we need independence to prove these facts? Write a couple sentences *in English* referencing the lecture where we went over this.

- (t) [harder] Let $Y = X + c$ where X is a r.v. either discrete or continuous and $c \in \mathbb{R}$ is a constant. Let $M_X(t)$ denote the mgf for the r.v. X . Find $M_Y(t)$, the mgf for the shifted r.v. Y as a function of $M_X(t)$.

- (u) [difficult] Let $Y = aX$ where X is a r.v. either discrete or continuous and $a \in \mathbb{R}$ is a constant. Let $M_X(t)$ denote the mgf for the r.v. X . Find $M_Y(t)$, the mgf for the scaled r.v. Y as a function of $M_X(t)$.

Problem 2

Here, we will be deriving mgf's of some of our brand name r.v.'s and using them to prove cool things about distribution theory.

- (a) [easy] Let $X \sim \text{Bernoulli}(p)$. Show that $\mathbb{E}[X^{37}] = p$ using the definition of expectation. I will begin the problem below for you:

$$\begin{aligned}\mathbb{E}[X^{37}] &= \sum_{x \in \text{Supp}[X]} x^{37} f(x) = \sum_{x=0}^1 x^{37} p^x (1-p)^{1-x} \\ &= \end{aligned}$$

- (b) [harder] Using the definition of the mgf, find $M_X(t)$, the mgf for a Bernoulli r.v.

- (c) [difficult] Show that $\mathbb{E}[X^{37}] = p$ using the mgf and the fact you proved in question 1(o). You may have to use English to explain what you're doing. I do not expect you to take 37 derivatives, but you should take at least two and see the pattern.

(d) [difficult] Let $T \sim \text{Binomial}(n, p)$. Find the mgf of X using the definition of the mgf. You will need to invoke the binomial theorem here (see class notes).

(e) [harder] Using the fact you proved in 1(r), use mgf's to show that $T = X_1 + \dots + X_n$ where T is the binomial r.v. from the previous problem and $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Bernoulli}(p)$.

(f) [difficult] Show that the sum of two independent r.v.'s $T_1 \sim \text{Binomial}(n_1, p)$ and $T_2 \sim \text{Binomial}(n_2, p)$ is a new r.v. which itself is a binomial and find its parameters. Does this make sense given what you proved in (e)?

(g) [easy] Let $X_1, \dots, X_r \stackrel{iid}{\sim} \text{Geometric}(p)$. We proved in class that the mgf for the geometric is:

$$M_{X_i}(t) = \frac{pe^t}{1 - (1-p)e^t}$$

Let $T = X_1 + \dots + X_r$. We know from our study of discrete r.v.'s that $T \sim \text{NegBin}(r, p)$ which was illustrated conceptually. However, this was never proven. We will prove it here. Using the fact that:

$$M_T(t) = \left(\frac{pe^t}{1 - (1-p)e^t} \right)^r$$

for the negative binomial r.v., show that negative binomial r.v.'s are indeed the sum of many $\stackrel{iid}{\sim}$ geometric r.v.'s using the fact you proved in 1(r). This looks hard but it is marked easy, so I assure you: it is easy! (It is not easy if you do not have mgf's, really really not easy).

- (h) [harder] Show that the sum of two independent r.v.'s $T_1 \sim \text{NegBin}(r_1, p)$ and $T_2 \sim \text{NegBin}(r_2, p)$ is itself a negative binomial and find its parameters.

- (i) [difficult] Let $X \sim \text{Poisson}(\lambda)$ with mgf:

$$M_X(t) = e^{\lambda(e^t - 1)}$$

Show that the sum of two independent r.v.'s $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$ is itself a Poisson r.v. and find its parameter. You don't even need to know what the Poisson is for this question.

- (j) [difficult] You have shown previously that the mgf for the r.v. $X \sim \text{Binomial}(n, p)$ was $M_X(t) = (1 - p + pe^t)^n$. In the previous problem you were given the mgf for the r.v. $X \sim \text{Poisson}(\lambda)$ was $M_X(t) = e^{\lambda(e^t - 1)}$. Although we didn't cover it in class, the Poisson distribution is the limit of the Binomial distribution with $n \rightarrow \infty$ and the rate parameter pinned $\lambda = np$. Prove that the Poisson's mgf is this limit of the Binomial's mgf. Really not hard: just some algebraic manipulations. Let $p = \lambda/n$.
- (k) [harder] We proved in class that if $Z \sim \mathcal{N}(0, 1)$ then $M_Z(t) = e^{t^2/2}$. We also showed that if $X = \sigma Z + \mu$, then $X \sim \mathcal{N}(\mu, \sigma^2)$. Show that the mgf of X is $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$ using what you learned in question 1 and the class notes.
- (l) [difficult] Let X_1, X_2, \dots, X_n be a sequence of independent normal random variables where X_1 has mean μ_1 and variance σ_1^2 and X_2 has mean μ_2 and variance σ_2^2 , etc. Show that $X_1 + X_2 + \dots + X_n$ is normally distributed and find its parameters.

(m) [E.C.] Let $X \sim \text{Exp}(\lambda)$ with mgf:

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

Demonstrate that the sum of $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Exp}(\lambda)$ is an Erlang r.v. with parameters n and λ , the continuous analogue of the Negative Binomial r.v. (You will need to look up the mgf of the Erlang distribution). Do on a separate piece of paper.

(n) [E.C.] The standard Cauchy distribution has center 0 and scale parameter 1 and its PDF is:

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

The Cauchy distribution is a classic example of a pathological r.v. You'll see why. The standard Cauchy r.v. is actually the ratio of two independent standard normal r.v.'s: if $X \sim \mathcal{N}(0, 1)$ and $Y \sim \mathcal{N}(0, 1)$ then X/Y has the PDF above.

Prove that the mgf for a standard Cauchy r.v. does not exist. This is really not hard but it looks menacing. Do on a separate piece of paper.

(o) [E.C.] Prove $\bar{X} \xrightarrow{\mathcal{D}} \mu$ which is a weak Law of Large Numbers. You will just need to copy from the class notes. Do on a separate piece of paper.

Central Limit Theorem We will introduce it here and provide some illustrations.

Problem 3

We will now prove the “Central Limit Theorem” (the CLT), the crown jewel of a college-level course on probability. I will not ask you to prove it on the exam, but I may ask you general questions about the proof. The best thing to do then, is to prove it yourself in which case you'll be prepared for basic questions. You will not be asked to do anything other than to critically read and synthesize the lecture notes.

(a) [harder] We start by looking at r.v.'s X_1, X_2, \dots, X_n . What are the assumptions on these r.v.'s? There are three.

(b) [easy] We started our proof examining \bar{X}_n . What is $\mathbb{E}[\bar{X}_n]$?

(c) [easy] What is $\text{SE}[\bar{X}_n]$. I'm marking this easy because you've done it before and it's in the notes.

(d) [easy] Below is its standardized form.

$$C_n = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

If it's "standardized," what is the mean and standard error of C_n ?

(e) [easy] Show that $C_n = \frac{Z_1}{\sqrt{n}} + \frac{Z_2}{\sqrt{n}} + \dots + \frac{Z_n}{\sqrt{n}}$.

(f) [harder] Find the first five terms of e^{tZ} using the Taylor Series. You can put $+$... after the five terms. This is the same thing you did in 1(g).

- (g) [easy] Find the first five terms of $M_Z(t)$, the mgf of all the Z_i 's (which are $\overset{iid}{\sim}$). This just means take the expectation of what you did in the previous problem since $M_Z(t) := \mathbb{E}[e^{tZ}]$. Remember: it's standardized! So you know something about the first two moments: $\mathbb{E}[Z]$ and $\mathbb{E}[Z^2]$. You can put $+$... after the terms. The five terms in the mgf should have become four (since one term vanished conveniently).

- (h) [easy] Using your answer to (h) and what you learned in 1(u), find the mgf of Z/\sqrt{n} and denote it $M_{\frac{Z}{\sqrt{n}}}(t)$. Make sure you remember your fractional powers e.g. $(\sqrt{n})^3 = n^{3/2}$. I've marked this easy since I give away the hard part of the answer in the next question.

- (i) [difficult] Argue that the tail of $M_{\frac{Z}{\sqrt{n}}}(t)$ which should look like

$$\text{tail} := \frac{t^3 \mathbb{E}[Z^3]}{3!n^{3/2}} + \frac{t^4 \mathbb{E}[Z^4]}{4!n^2} + \dots$$

belongs to $o(1/n)$ which means that

$$\lim_{n \rightarrow \infty} \frac{\text{tail}}{\frac{1}{n}} = 0$$

which *in English* means that the tail drops to zero *more quickly* than $\frac{1}{n}$. Remember, when you take a limit such as $\lim_{n \rightarrow \infty}$, the t 's and the $\mathbb{E}[Z^k]$ terms are constants.

(j) [easy] Write $M_{\frac{Z}{\sqrt{n}}}(t)$ using “little-o” notation using what you showed in the previous problem.

(k) [easy] Find the mgf of C_n denoted $M_{C_n}(t)$ using the fact that you have the sum n independent $\frac{Z_i}{\sqrt{n}}$ r.v.’s. Hint: answer looks something like what’s in the next question.

(l) [E.C.] Prove for any $c \in \mathbb{R}$ that

$$\lim_{n \rightarrow \infty} \left(1 + \frac{c}{n} + o\left(\frac{1}{n}\right) \right)^n = e^c$$

that is, show that additive terms that drop to zero more quickly than $1/n$ in the limit do not matter so they can be effectively ignored to arrive at the familiar limit:

$$\lim_{n \rightarrow \infty} \left(1 + \frac{c}{n} \right)^n = e^c$$

I’m sure the answer is on Google somewhere! Do on a separate piece of paper.

(m) [easy] Regardless of whether or not you proved the previous extra credit, use it **as fact** and show that the limiting mgf of C ,

$$M_C(t) := \lim_{n \rightarrow \infty} M_{C_n}(t)$$

has the same form as a standard normal mgf. This mgf is found in problem 2(1). Thus, since it has the same fingerprint, it must have the standard normal distribution! QED.

- (n) [easy] Congratulations: you have proved the most basic central limit theorem, the crown jewel of Math 241! Write a smiley face by the crown.



Problem 4

We've now proven the CLT, but you may not viscerally feel it yet since all you did is just a bunch of math in bilateral Laplace transform space (which doesn't exactly exalt the spirit). This problem is here to make you really understand the power of this technology. This is a preview to confidence intervals and hypothesis testing which will be the topics of the remainder of this course.

- (a) [easy] Image $X \sim \text{Bernoulli}(\frac{1}{2})$. Run the following code in R and print out (in black and white) the result of the following. Does this look like the PMF of X ?

```
#first line placeholder
par(mfrow = c(1, 1))
N = 10000
xs = rbinom(N, 1, 0.5)
h = hist(xs, breaks = 100, plot = FALSE)
h$counts = h$counts / sum(h$counts)
plot(h)
#last line placeholder
```

- (b) [easy] Now we're going to imagine $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Bernoulli}(p = 0.5)$ and compute our friend:

$$C_n = \sqrt{n}\bar{Z} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Since we know $\mu = p = 0.5$ and $\sigma = \sqrt{p(1-p)} = \sqrt{0.5(1-0.5)} = 0.5$,

$$C_n = \frac{\bar{X} - 0.5}{\frac{0.5}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient):

This will look at the estimated PDF for C_n for $n = 10, 100, 1000, 5000, 10000, 50000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."

```
#first line placeholder
par(mfrow = c(3, 2))
N = 10000
mu = 0.5
sigma = 0.5
for (n in c(10, 100, 1000, 5000, 10000, 50000)){
  xs = matrix(rbinom(n * N, 1, mu), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder
```

- (c) [easy] Now we're going to imagine $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Bernoulli}(p = 0.1)$ which is more skewed towards failures. Since we know $\mu = p = 0.1$ and $\sigma = \sqrt{p(1-p)} = \sqrt{0.1(1-0.1)} = 0.3$,

$$C_n = \frac{\bar{X} - 0.1}{\frac{0.3}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient).

This will look at the estimated PDF for C_n for $n = 10, 100, 1000, 5000, 10000, 50000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."

```
#first line placeholder
par(mfrow = c(3, 2))
N = 10000
mu = 0.1
sigma = 0.3
for (n in c(10, 100, 1000, 5000, 10000, 50000)){
  xs = matrix(rbinom(n * N, 1, mu), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder
```

- (d) [easy] Now we're going to imagine $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Binomial}(n = 100, p = 0.5)$. Since we know $\mu = np = 50$ and $\sigma = \sqrt{np(1-p)} = \sqrt{100(0.5)(1-0.5)} = 5$,

$$C_n = \frac{\bar{X} - 50}{\frac{5}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient).

This will look at the estimated PDF for C_n for $n = 10, 100, 1000, 5000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."


```

#first line placeholder
par(mfrow = c(2, 2))
N = 10000
mu = 50
sigma = 5
for (n in c(10, 100, 1000, 5000)){
  xs = matrix(rbinom(n * N, 100, 0.5), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder

```

- (e) [easy] Now we're going to imagine $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Poisson}(\lambda = 7)$. Since we know $\mu = \lambda = 7$ and $\sigma = \sqrt{\lambda} = \sqrt{7} = 2.6458$,

$$C_n = \frac{\bar{X} - 7}{\frac{2.6458}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient).

This will look at the estimated PDF for C_n for $n = 10, 100, 1000, 5000, 10000, 50000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."

```

#first line placeholder
par(mfrow = c(3, 2))
N = 10000
mu = 7
sigma = 2.6458
for (n in c(10, 100, 1000, 5000, 10000, 50000)){
  xs = matrix(rpois(n * N, mu), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder

```

- (f) [easy] Now we're going to imagine $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Exp}(\lambda = 7)$. Since we know $\mu = 1/\lambda = 0.1429$ and $\sigma = 1/\lambda = 0.1429$,

$$C_n = \frac{\bar{X} - 0.1429}{\frac{0.1429}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient).

This will look at the estimated PDF for C_n for $n = 10, 100, 1000, 5000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."

```

#first line placeholder
par(mfrow = c(2, 2))
N = 10000
mu = 0.1429
sigma = 0.1429
for (n in c(10, 100, 1000, 5000)){
  xs = matrix(rexp(n * N, 7), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder

```

- (g) [easy] Now we're going to imagine $X_1, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu = 3, \sigma^2 = 6^2)$. Since we know $\mu = 3$ and $\sigma = 6$,

$$C_n = \frac{\bar{X} - 3}{\frac{6}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient).

This will look at the estimated PDF for C_n for $n = 10, 100, 1000, 5000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."

```
#first line placeholder
par(mfrow = c(2, 2))
N = 10000
mu = 3
sigma = 6
for (n in c(10, 100, 1000, 5000)){
  xs = matrix(rnorm(n * N, 3, 6), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder
```

- (h) [harder] Why did this converge immediately? See question 2(m) for the answer.

- (i) [easy] Now we're going to get more elaborate. Imagine the following PDF:

I will consider this to be called the "bathtub function." This is not a real brand name distribution. I just made it up and it won't be on the test. Write about why you think I called it the "bathtub function." Does this look like a bell curve at all??

```

#first line placeholder
par(mfrow = c(1, 1))
N = 10000
xs = 100 * rbeta(N, 0.1, 0.1)
h = hist(xs, breaks = 1000, plot = FALSE)
h$counts = h$counts / sum(h$counts)
plot(h)
#last line placeholder

```

- (j) [easy] Now we're going to make sure the central limit theorem works even with stuff that's as crazy as the bathtub. Imagine $X_1, \dots, X_n \stackrel{iid}{\sim}$ bathtub. By advanced math, I know $\mu = 50$ and $\sigma = 45.6436$ thus,

$$C_n = \frac{\bar{X} - 50}{\frac{45.6436}{\sqrt{n}}}$$

We are going to look at the distribution of C_n for different values of n . Remember $C = \lim_{n \rightarrow \infty} C_n$ should be distributed as $\mathcal{N}(0, 1)$. Run the following code in R and print out (in black and white) the result (please be patient).

This will look at the estimated PDF for C_n for $n = 2, 5, 10, 50, 100, 1000$. At what n value does it appear that C_n converged to a $\mathcal{N}(0, 1)$? Pay attention to gaps of whitespace in these plots. They mean there's no support there! The standard normal has support everywhere. If it never converges, write "never."

```

#first line placeholder
par(mfrow = c(3, 2))
N = 10000
mu = 50
sigma = 45.6436
for (n in c(2, 5, 10, 50, 100, 1000)){
  xs = matrix(100 * rbeta(n * N, 0.1, 0.1), ncol = N)
  xbars = (colMeans(xs) - mu) / (sigma / sqrt(n))
  hist(xbars, breaks = 100, xlim = c(-4, 4),
       main = paste("PDF estimate of Cn for n =", n), col = "blue")
}
#last line placeholder

```

- (k) [easy] Let's sum up what we've learned in this problem. The central limit states as $n \rightarrow \infty$ (i.e. it gets big) then C_n becomes a standard normal. Does it become a standard normal at different rates depending on the distribution of the r.v. being sampled? Yes/no is fine.

Problem 5

Now that we've now proven the CLT, we are going to use it a little bit. Assume $X_1, \dots, X_n \stackrel{iid}{\sim}$ something with mean μ and standard error σ .

- (a) [easy] Assume n is large enough for the CLT to kick in, show that:

$$\bar{X} \sim \mathcal{N}\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$$

- (b) [easy] Assume n is large enough for the CLT to kick in, show that:

$$T_n \sim \mathcal{N}\left(n\mu, (\sqrt{n}\sigma)^2\right)$$

- (c) [easy] Let $X_1, \dots, X_n \stackrel{iid}{\sim}$ Bernoulli(p) and let $\hat{P} := \bar{X}$. Assume n is large enough for the CLT to kick in, show that:

$$\hat{P} \sim \mathcal{N}\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$$