1. The lifetime of a certain brand of industrial light bulb, Y (in months), follows an exponential distribution with mean $\beta = 18$. An iid sample of lifetimes $Y_1, Y_2, ..., Y_{36}$ corresponding to $n = 36$ light bulbs will be observed.
(a) Find the exact sampling distribution of $\overline{Y}$.
(b) Find the approximate sampling distribution of $\overline{Y}$ by using the CLT.
(c) Use your answer in part (b) to approximate $P(\overline{Y} > 21)$.

2. Suppose $Z_1, Z_2, ..., Z_6$ is an iid sample from a $\mathcal{N}(0,1)$ distribution. Suppose $Z_7 \sim \mathcal{N}(0,1)$ and that Z_7 is independent of $Z_1, Z_2, ..., Z_6$. Find the sampling distribution of

(a) $\overline{Z} = \frac{1}{6}\sum_{i=1}^{6} Z_i$

(b) $T = \sum_{i=1}^{6}(Z_i - \overline{Z})^2$

(c) $U = \sqrt{3}Z_7/\sqrt{Z_1^2 + Z_2^2 + Z_3^2}$

(d) $V = (Z_1^2 + Z_2^2 + Z_3^2)/(Z_4^2 + Z_5^2 + Z_6^2)$.

3. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample of Bernoulli observations with mean p, where $0 < p < 1$. Recall the sample proportion is

$$\widehat{p} = \frac{X}{n},$$

where $X = Y_1 + Y_2 + \cdots + Y_n$. Prove that $E(\widehat{p}) = p$ and that

$$V(\widehat{p}) = \frac{p(1-p)}{n}.$$

(c) For a fixed sample size, what value of p makes $V(\widehat{p})$ as large as possible?

4. Suppose $U \sim \chi^2(64)$; i.e., U has a χ^2 distribution with $\nu = 64$ degrees of freedom.
(a) Use the Central Limit Theorem (CLT) to approximate $P(60 < U < 70)$. State clearly how you are using the CLT.
(b) If $V \sim \chi^2(8)$, independent of U, find a function of U and V that has an F distribution. What are the degrees of freedom associated with your distribution?

5. Suppose $Y_1, Y_2, ..., Y_5$ is an iid sample of size $n = 5$ from a $\mathcal{N}(0,1)$ distribution. Let $\overline{Y}$ and S^2 denote the sample mean and sample variance, respectively. Define the statistics

$$\begin{aligned} T_1 &= \frac{Y_1^2}{(Y_2^2 + Y_3^2)/2} \\ T_2 &= \frac{\overline{Y}}{\sqrt{S^2/5}} \\ T_3 &= 4S^2 + (\sqrt{5}\overline{Y})^2. \end{aligned}$$

Give the precise sampling distribution of each statistic. You need not derive anything here rigorously; you can use results we discussed/proved in class.

6. A professor who teaches mathematical statistics observes

$$Y = \text{ the number of absent students}$$

each day that class meets. There are $n = 25$ days of class throughout the semester, so he observes $Y_1, Y_2, ..., Y_{25}$. Treat these observations as an iid sample from a Poisson distribution with mean $\lambda = 4$.
(a) Derive the moment generating function of

$$U = Y_1 + Y_2 + \cdots + Y_{25},$$

the total number of absences he observes throughout the semester. What is the exact sampling distribution of U?
(b) Use the CLT to determine the approximate sampling distribution of U.
(c) Use your answer in part (b) to approximate the probability U will be greater than or equal 125 days.

7. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample of exponential($\beta = 1$) random variables. Define the statistics

$$\begin{aligned} U_n &= 2(Y_1 + Y_2 + \cdots + Y_n) \\ V_n &= \sqrt{n}(\overline{Y} - 1). \end{aligned}$$

(a) Derive the moment generating function (mgf) of U_n. What is the distribution of U_n?
(b) Show the mgf of V_n, for $t < \sqrt{n}$, is

$$m_{V_n}(t) = \{e^{t/\sqrt{n}} - (t/\sqrt{n})e^{t/\sqrt{n}}\}^{-n}.$$

(c) For any $t < \sqrt{n}$, find

$$\lim_{n \to \infty} m_{V_n}(t).$$

Hint: Think about the CLT when you are examining the V_n sequence.

8. I have four statistics T_1, T_2, T_3, and T_4. I have determined

- $T_1 \sim \mathcal{N}(0, 1)$
- $T_2 \sim \mathcal{N}(-3, 4)$
- $T_3 \sim \chi^2(3)$
- $T_4 \sim \chi^2(5)$
- T_1, T_2, T_3, and T_4 are mutually independent.

(a) What is the distribution of $T_1 - T_2$?
(b) Find a function of T_1, T_2, T_3, and T_4 that has a t distribution with 8 degrees of freedom.
(c) Find a function of T_1, T_2, T_3, and T_4 that has an F distribution with 4 (numerator) and 6 (denominator) degrees of freedom.

9. Let Y denote the time (in minutes) it takes for a customer representative to respond to a telephone inquiry. It is assumed Y follows a uniform distribution from 0.5 to 3.5. In a 2-hour shift, a typical representative will respond to $n = 15$ calls with response times $Y_1, Y_2, ..., Y_{15}$. Treating these times as an iid sample, approximate the probability the sample mean response time $\overline{Y}$ will be less than 2.15 minutes.

10. Body mass index (BMI) is commonly used to estimate a healthy body weight based on how tall a person is. Suppose a random sample of $n = 26$ severely obese patients is observed and that we model the BMI measurements $Y_1, Y_2, ..., Y_{26}$ as an iid sample from a $\mathcal{N}(40, 25)$ distribution. Let $\overline{Y}$ and S^2 denote the sample mean and sample variance, respectively.
(a) What is the sampling distribution of $\overline{Y}$?
(b) What is the sampling distribution of S^2?
(c) Describe how you would find the joint probability $P(\overline{Y} > 41,\ S^2 < 16.5)$.

11. Suppose $Y_1, Y_2, ..., Y_{36}$ is an iid sample of size $n = 36$ from a geometric population distribution with $p = 0.75$.
(a) Derive the moment generating function of

$$T = Y_1 + Y_2 + \cdots + Y_{36}.$$

Does T have a distribution that you recognize? If so, which one?
(b) Use the Central Limit Theorem to approximate $P(1.25 < \overline{Y} < 1.50)$.

12. Suppose

- $Y \sim \mathcal{N}(1, 4)$; i.e., Y has a normal distribution with $\mu = 1$ and $\sigma^2 = 4$
- $U \sim \chi^2(6)$
- $V \sim F(6, 1)$
- Y and U are independent.

(a) Find a function of Y and U that has a t distribution with 6 degrees of freedom.
(b) Find a function of Y and U that has a χ^2 distribution with 7 degrees of freedom.
(c) Find a function of Y and U that has the same distribution as $1/V$.

13. Suppose $Y_1, Y_2, ..., Y_n$ is an iid $\mathcal{N}(\mu, \sigma^2)$ sample. Let $\overline{Y}$ and S^2 denote the sample mean and sample variance, respectively.
(a) Find the moment generating function of $\overline{Y}$.
(b) Find the moment generating function of S^2.
(c) Find a function of $\overline{Y}$ and S^2 that has a t distribution.

14. Suppose we have two independent samples:

$$Y_{11}, Y_{12}, ..., Y_{1n} \sim \text{iid } \mathcal{N}(\mu_1, \sigma^2)$$
$$Y_{21}, Y_{22}, ..., Y_{2n} \sim \text{iid } \mathcal{N}(\mu_2, \sigma^2).$$

Notice the two population variances are the same and that the sample sizes are equal. The population means μ_1 and μ_2 are unknown parameters. The common population variance σ^2 is also an unknown parameter. Define

$$
\begin{aligned}
\overline{Y}_{1+} = \frac{1}{n}\sum_{j=1}^{n} Y_{1j} &= \text{sample mean for sample 1} \\
\overline{Y}_{2+} = \frac{1}{n}\sum_{j=1}^{n} Y_{2j} &= \text{sample mean for sample 2} \\
S_1^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_{1j} - \overline{Y}_{1+})^2 &= \text{sample variance for sample 1} \\
S_2^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_{2j} - \overline{Y}_{2+})^2 &= \text{sample variance for sample 2.}
\end{aligned}
$$

(a) Find the sampling distribution of $U = \overline{Y}_{1+} - \overline{Y}_{2+}$.
(b) Find the distribution of

$$W = \frac{(n-1)S_1^2}{\sigma^2} + \frac{(n-1)S_2^2}{\sigma^2} = \frac{(n-1)S_1^2 + (n-1)S_2^2}{\sigma^2}.$$

(c) Explain why U and W are independent.
(d) If $\mu_1 - \mu_2 = 0$, find a statistic, which is a function of U and W, that has an F distribution. What are its degrees of freedom? Remember a statistic can not depend on unknown parameters.

15. Let $Y_1, Y_2, ..., Y_{953}$ denote CD4 count measurements for $n = 953$ Senegalese sex workers. Assume $Y_1, Y_2, ..., Y_{953}$ are iid gamma random variables with shape $\alpha = 9$ and scale $\beta = 130$.
(a) Use the CLT to approximate the probability the sample mean $\overline{Y}$ will be between 1150 and 1200.
(b) Your answer in part (a) is based on an approximate sampling distribution. Explain how you could compute $P(1150 < \overline{Y} < 1200)$ exactly without using an approximation.

16. Suppose $Y_1, Y_2, ..., Y_n$ is an iid $\mathcal{N}(0, \sigma^2)$ sample. Let $\overline{Y}$ and S^2 denote the sample mean and sample variance, respectively.
(a) State the sampling distribution of $\overline{Y}$.
(b) Find $V(S^2)$.
(c) Define

$$Q_1 = \frac{1}{\sigma^2}\left[n\overline{Y}^2 + (n-1)S^2\right] \quad \text{and} \quad Q_2 = \frac{n\overline{Y}^2}{S^2}.$$

Argue that $Q_1 \sim \chi^2(n)$ and that $Q_2 \sim F(1, n-1)$. These arguments need not be overly mathematical, but they must be convincing.

17. In an environmental study, Y denotes the proportion of particulate matter that is deemed hazardous. Empirical evidence from data collected by the EPA reveals the probability density function (pdf) of Y is

$$f_Y(y) = \begin{cases} 4(1-y)^3, & 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) A total of $n = 36$ particulate matter measurements will be taken, yielding the random sample $Y_1, Y_2, ..., Y_{36}$. Treating these measurements as an iid sample from $f_Y(y)$, use the Central Limit Theorem to approximate $P(\overline{Y} < 0.175)$.
(b) Find the smallest sample size n so that $P(\overline{Y} > 0.22) \approx 0.01$.

18. I have 4 statistics T_1, T_2, T_3, and T_4. I know that

- T_1, T_2, T_3, and T_4 are (mutually) independent
- $T_1 \sim \chi^2(4)$
- $T_2 \sim \chi^2(5)$
- $T_3 \sim \mathcal{N}(1, 4)$
- T_4 has the same distribution as $T_1 + T_2$.

(a) Show that the moment generating function of $Z = (T_3 - 1)/2$ is $m_Z(t) = \exp(t^2/2)$. What is the distribution of Z?
(b) Find a statistic that has a $\chi^2(1)$ distribution.
(c) Find a statistic that has an $F(9, 4)$ distribution.
(d) Find a statistic that has a $t(18)$ distribution.
(e) Find a statistic that has a gamma distribution with shape parameter $\alpha = 2$ and scale parameter $\beta = 6$.

19. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a $\mathcal{N}(\mu, \sigma^2)$ distribution. Let $\overline{Y}$ and S^2 denote the sample mean and sample variance, respectively.
(a) Find the distribution of

$$Q = \frac{(\overline{Y} - \mu)^2}{\sigma^2/n} + \frac{(n-1)S^2}{\sigma^2}.$$

(b) When $\mu = 0$, formulate an approach to determine the value of c that satisfies

$$P\left(-c < \frac{S}{\overline{Y}} < c\right) = 1 - \alpha,$$

for $\alpha \in (0, 1)$. The statistic $S/\overline{Y}$ is called the *coefficient of variation*; this is a measure of variation relative to the mean.

20. Suppose $X_1, X_2, ..., X_n$ is an iid sample from a $\mathcal{N}(0, \sigma^2)$ distribution, where $\sigma^2 > 0$. Define the statistics

$$T_1 = \left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \quad \text{and} \quad T_2 = \frac{1}{n}\sum_{i=1}^{n} |X_i|.$$

(a) Derive the sampling distribution of T_1.
(b) Calculate $E(T_1)$ and $E(T_2)$ and establish an inequality between them.
(c) Calculate $V(T_2)$.

21. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a $\mathcal{N}(\mu, \sigma^2)$ population distribution. Let $\overline{Y}$ and S^2 denote the usual sample mean and sample variance, respectively. Let

$$\begin{aligned} Q_1 &= \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} \\ Q_2 &= \frac{(n-1)S^2}{\sigma^2}. \end{aligned}$$

(a) What is the distribution of Q_1? the distribution of Q_2?
(b) Determine the distribution of $Q_1^2 + Q_2$.
(c) Suppose Y_{n+1} is a "new observation" that also follows a $\mathcal{N}(\mu, \sigma^2)$ distribution and suppose Y_{n+1} is independent of $Y_1, Y_2, ..., Y_n$. Argue that

$$\frac{Y_{n+1} - \overline{Y}}{S\sqrt{1 + \frac{1}{n}}} \sim t(n-1),$$

where S is the sample standard deviation.

22. The hospitalization period, in days, for patients following treatment for a serious kidney disorder is a random variable $U = Y + 2$, where Y has pdf

$$f_Y(y) = \begin{cases} \dfrac{24}{(y+2)^4}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine $f_U(u)$, the pdf of U. Make sure to note the support.
(b) Suppose a random sample of $n = 20$ patients with this disorder is available and each patient's hospitalization period will be observed, producing the iid random variables $U_1, U_2, ..., U_{20}$ from $f_U(u)$. Use the CLT to approximate

$$P(\overline{U} > 5),$$

where $\overline{U}$ is the sample mean of $U_1, U_2, ..., U_{20}$.

23. Suppose Y_1, Y_2 is an independent and identically distributed (iid) sample of size $n = 2$ from a $\mathcal{U}(0, 1)$ population distribution; i.e., the population probability density function (pdf) is

$$f_Y(y) = \begin{cases} 1, & 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Use the cumulative distribution function (cdf) technique to show the pdf of the sample sum $T = Y_1 + Y_2$ is given by

$$f_T(t) = \begin{cases} t, & 0 < t \le 1 \\ 2 - t, & 1 < t < 2 \\ 0, & \text{otherwise.} \end{cases}$$

Why do you think $f_T(t)$ is called a "triangular density?"

(b) The (sampling) distribution you derived in part (a) is an "exact" result. That is, $f_T(t)$ above describes the true sampling distribution of T.

1. What is the approximate sampling distribution of T conferred by the Central Limit Theorem (CLT)? Note that $n = 2$. How well do you think the CLT "works" in this case?
2. Calculate $P(|T - 1| > 0.5)$ exactly using $f_T(t)$. Then approximate $P(|T - 1| > 0.5)$ by using the CLT. Comment on the differences in these answers.

24. There are approximately 540 coronavirus testing locations in South Carolina. At the beginning of the day, suppose officials at each testing location record

$$Y = \text{number of specimens tested to find the first positive case,}$$

and assume Y follows a geometric distribution with probability of success p.
(a) For purposes of data analysis, suppose we assume $Y_1, Y_2, ..., Y_{540}$ are independent and identically distribution (iid) and arise from a geometric(p) population-level model. Provide 2-3 practical reasons why these assumptions might be incorrect (think about what you know about coronavirus).
(b) Suppose $Y_1, Y_2, ..., Y_{540}$ are iid from a geometric(p) population distribution. Derive the moment generating function (mgf) of

$$T = \sum_{i=1}^{540} Y_i.$$

What is the exact sampling distribution of T? By "exact," I mean do not appeal to the Central Limit Theorem (CLT).
(c) Suppose $Y_1, Y_2, ..., Y_{540}$ are iid from a geometric(p) population distribution. If $p = 0.20$, use the CLT to approximate the probability the sample sum T will be larger than 3000.

25. *Fun with the t distribution.* Suppose the random variable Y follows a t distribution with $\nu > 0$ degrees of freedom.
(a) State the definition of a $t(\nu)$ random variable in terms of a standard normal random variable and a χ^2 random variable. Be precise.
(b) Recall the *skewness* of a random variable Y is given by

$$\xi = \frac{E[(Y - \mu)^3]}{\sigma^3},$$

where $\mu = E(Y)$ and $\sigma^2 = V(Y)$. Derive ξ above when $Y \sim t(\nu)$. What conditions are needed on ν for ξ to exist? Prove any claims you make.
(c) When $Y_1, Y_2, ..., Y_n$ form an iid sample from a $\mathcal{N}(\mu, \sigma^2)$ population distribution, we showed in class that

$$T = \frac{\overline{Y} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

where $\overline{Y}$ and S are the sample mean and sample standard deviation, respectively. Using this as a starting point, write out what T^2 is and then show T^2 can be written as the ratio of two independent χ^2 random variables, each divided by their respective degrees of freedom.

26. Actuaries have determined the insurance claim amount Y for a group of policy holders has probability density function (pdf)

$$f_Y(y) = \begin{cases} \dfrac{24y(10+2y-0.3y^2)}{10000}, & 0 < y < 10 \\ 0, & \text{otherwise.} \end{cases}$$

An auditor selects a random sample (iid sample) of $n = 40$ claims and records Y for each one, producing $Y_1, Y_2, ..., Y_{40}$.
(a) Approximate the probability the sample sum $T = \sum_{i=1}^{40} Y_i$ is between 200 and 250.
(b) Approximate the probability the sample mean $\overline{Y}$ is less than 5.5.
(c) Let S^2 denote the sample variance of $Y_1, Y_2, ..., Y_{40}$. Are the sample mean $\overline{Y}$ and S^2 independent? Explain.
(d) Why can't we compute the probabilities in parts (a) and (b) exactly?

27. *Revisiting the CLT proof.* Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a population distribution with mean $E(Y_i) = \mu$ and variance $V(Y_i) = \sigma^2 < \infty$, and define

$$Z_n = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}}.$$

In the notes, we argued $m_{Z_n}(t) = [m_U(t/\sqrt{n})]^n$, where $m_U(\cdot)$ is the moment generating function (mgf) of

$$U_i = \frac{Y_i - \mu}{\sigma},$$

for $i = 1, 2, ..., n$. To finish the proof of the CLT, we wrote $m_U(t/\sqrt{n})$ in its McLaurin series expansion and showed $m_{Z_n}(t) \to e^{t^2/2}$, as $n \to \infty$. An alternative proof of the CLT establishes $m_{Z_n}(t) = [m_U(t/\sqrt{n})]^n$ but then takes a different approach going forward. This different approach is outlined below in three parts.
(a) Using the properties of logarithms and L'Hôpital's Rule, show

$$\lim_{n\to\infty} \ln\left([m_U(t/\sqrt{n})]^n\right) = \frac{t}{2}\left[\lim_{n\to\infty} \frac{m'_U(t/\sqrt{n})/m_U(t/\sqrt{n})}{1/\sqrt{n}}\right].$$

(b) Argue $m'_U(0) = 0$ and thus L'Hôpital's Rule can be applied again. Doing so, show the expression in part (a) on the right-hand side is equal to

$$\frac{t^2}{2}\left\{\lim_{n\to\infty} \frac{m''_U(t/\sqrt{n})m_U(t/\sqrt{n}) - [m'_U(t/\sqrt{n})]^2}{[m_U(t/\sqrt{n})]^2}\right\}.$$

(c) Show $m''_U(t/\sqrt{n}) \to 1$ as $n \to \infty$ and hence argue the limit directly above is equal to 1. This allows you to conclude

$$\lim_{n\to\infty} \ln\left([m_U(t/\sqrt{n})]^n\right) = \frac{t^2}{2}$$

after two applications of L'Hôpital's Rule. Finally, argue

$$\lim_{n\to\infty} m_{Z_n}(t) = \exp\left\{\lim_{n\to\infty} \ln\left([m_U(t/\sqrt{n})]^n\right)\right\}$$

and you are done.

28. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a population distribution with mean $E(Y) = \mu$, variance $V(Y) = \sigma^2$, and $E(Y^4) < \infty$; i.e., the fourth population moment is finite, and let S^2 denote the sample variance. A general expression for $V(S^2)$ is

$$V(S^2) = \frac{1}{n}\left[\mu_4 - \left(\frac{n-3}{n-1}\right)\sigma^4\right],$$

where

$$\mu_4 = E[(Y-\mu)^4]$$

is the fourth central moment.
(a) By "general expression," I mean the formula above for $V(S^2)$ applies to *any* population distribution with $E(Y^4) < \infty$. When the population distribution is $\mathcal{N}(\mu, \sigma^2)$, show the formula for $V(S^2)$ reduces to

$$V(S^2) = \frac{2\sigma^4}{n-1}.$$

(b) The quantity $V(S^2)$ involves two symbols that mean "variance," the V and S^2. What is the difference? Explain specifically what $V(S^2)$ measures.
(c) The fourth central moment μ_4 also appears in the formula for the population kurtosis; i.e.,

$$\kappa = \frac{E[(Y-\mu)^4]}{(\sigma^2)^2},$$

where μ is the population mean and σ^2 is the population variance. When the population distribution is $\mathcal{N}(\mu, \sigma^2)$, what does the population kurtosis equal?

29. Particles are subject to collisions that cause them to split into two parts with each part a fraction of the parent. Suppose this fraction Y for any one split is uniformly distributed over $(0, 1)$. Following a single particle through n mutually independent splits, we obtain a fraction of the original particle

$$T = \prod_{i=1}^{n} Y_i.$$

Our goal in this problem is to derive the sampling distribution of T, when $Y_1, Y_2, ..., Y_n$ are regarded as iid $\mathcal{U}(0, 1)$. We will do this in the following two parts.
(a) Show that

$$U = -\ln T = \sum_{i=1}^{n} -\ln Y_i$$

and argue the statistic U has a gamma sampling distribution with shape parameter $\alpha = n$ and scale parameter $\beta = 1$.
(b) Because $U \sim \text{gamma}(n, 1)$, we can derive the sampling distribution of T by noting

$$T = h(U) = e^{-U}.$$

Carry out a transformation argument to show the probability density function (pdf) of T is

$$f_T(t) = \begin{cases} \dfrac{1}{\Gamma(n)}(-\ln t)^{n-1}, & 0 < t < 1 \\ 0, & \text{otherwise.} \end{cases}$$

(c) Calculate $P(T > 0.1)$ when $n = 5$.

30. I have five mutually independent statistics T_1, T_2, T_3, T_4, and T_5. I have determined these statistics have the following sampling distributions:

- $T_1 \sim \chi^2(3)$
- $T_2 \sim \chi^2(8)$
- $T_3 \sim \mathcal{N}(0,1)$
- $T_4 \sim \mathcal{N}(2,9)$, where $V(T_4) = 9$.
- $T_5 \sim F(3,8)$

On each part below, find a function of T_1, T_2, T_3, T_4, and T_5 that has the stated distribution. You don't have to use all five statistics on each part. You might only use one, two, or three of them. You don't need to be overly mathematical, but you do need to make a convincing argument your answer is correct.
(a) $\chi^2(11)$ distribution
(b) $t(3)$ distribution
(c) $\mathcal{N}(0,4)$ distribution
(d) $\mathcal{N}(2,8)$ distribution
(e) $F(8,3)$ distribution
(f) gamma distribution with $\alpha = 2$ and $\beta = 4$
(g) standard Cauchy distribution
(h) $F(1,11)$ distribution

31. An actuary models the lifetime of a home protection device (in years) using the random variable $Y = 3X^{0.5}$, where X is an exponential random variable with mean 1.
(a) Find the probability density function (pdf) of Y. Verify your pdf integrates to 1 over the support of Y.
(b) The actuary will observe a random sample of $n = 25$ homes in her area, producing the iid device lifetimes $Y_1, Y_2, ..., Y_{25}$. What is the (approximate) probability the sample mean $\overline{Y}$ exceeds 3 years?

32. Suppose Y is a random variable with mean $E(Y) = \mu$, variance $V(Y) = \sigma^2$, and moment generating function $m_Y(t)$. Assume $m_Y(t)$ exists for all t in an open neighborhood about zero.
(a) Show that $m_Y(0) = 1$.
(b) Show that

$$\lim_{t \to 0} \frac{\ln m_Y(t) - \mu t}{t^2} = \frac{\sigma^2}{2}.$$

(c) Now suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a population distribution with mean $E(Y) = \mu$ and variance $V(Y) = \sigma^2 < \infty$. Define

$$Z_n = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}}.$$

Show that $E(Z_n) = 0$, $V(Z_n) = 1$, and calculate

$$\lim_{n \to \infty} \frac{\ln m_{Z_n}(t)}{t^2}.$$