Law of Large Numbers and Moment Generating Functions

Matthew Seguin

1.

Let $S_n \sim \text{Binomial}(n, p)$.

We are asked to calculate the following for the cases $n = 100$ and $p_i = \frac{i}{10}$ for $i \in \{1, 2, ..., 10\}$ and $\epsilon = \frac{1}{10}$:

$$\mathbb{P}[\frac{S_n}{n} \geq p_i + \epsilon]$$

a.

First we will compute the exact probabilities. Recall that $\mathbb{P}[S_n = k] = \binom{n}{k} p^k (1-p)^{n-k}$ for $k \in \{0, 1, 2, ..., n\}$.

Therefore:

$$\mathbb{P}[\frac{S_n}{n} \geq p + \epsilon] = \mathbb{P}[S_n \geq n(p+\epsilon)] = \sum_{n(p+\epsilon) \leq k \leq n} \mathbb{P}[S_n = k] = \sum_{n(p+\epsilon) \leq k \leq n} \binom{n}{k} p^k (1-p)^{n-k}$$

If we let $n = 100$, $p_i = \frac{i}{10}$, and $\epsilon = \frac{1}{10}$ this reduces to:

$$\mathbb{P}[\frac{S_n}{n} \geq p_i + \epsilon] = \sum_{n(p_i+\epsilon) \leq k \leq n} \binom{n}{k} p_i^k (1-p_i)^{n-k} = \sum_{100(\frac{i}{10}+\frac{1}{10}) \leq k \leq 100} \binom{100}{k} \left(\frac{i}{10}\right)^k \left(1 - \frac{i}{10}\right)^{100-k}$$

$$= \sum_{10(i+1) \leq k \leq 100} \binom{100}{k} \left(\frac{i}{10}\right)^k \left(1 - \frac{i}{10}\right)^{100-k}$$

From the code shown **here** we got the probabilities below:

i	$\mathbb{P}[S_n/n \geq p_i + \epsilon]$
1	0.00198
2	0.0112
3	0.021
4	0.0271
5	0.0284
6	0.0248
7	0.0165
8	0.0057
9	0.0000266

b.

Now we will compute the Markov upper bound for these probabilities. Recall that $\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$ for $a > 0$ and a non-negative random variable X. Further recall that since $S_n \sim \text{Binomial}(n, p)$ we know $\mathbb{E}[S_n] = np$.

Therefore:

$$\mathbb{P}[\frac{S_n}{n} \geq p + \epsilon] = \mathbb{P}[S_n \geq n(p + \epsilon)] \leq \frac{\mathbb{E}[S_n]}{n(p + \epsilon)} = \frac{np}{n(p + \epsilon)} = \frac{p}{p + \epsilon}$$

If we let $n = 100$, $p_i = \frac{i}{10}$, and $\epsilon = \frac{1}{10}$ this reduces to:

$$\mathbb{P}[\frac{S_n}{n} \geq p_i + \epsilon] \leq \frac{p_i}{p_i + \epsilon} = \frac{i/10}{i/10 + 1/10} = \frac{i}{i + 1}$$

From the code shown **here** we got the probabilities bounds below:

i	Markov bound: $i/(i+1)$
1	0.50
2	0.667
3	0.75
4	0.80
5	0.833
6	0.857
7	0.875
8	0.889
9	0.90

c.

Now we will compute the Chebyschev upper bound for these probabilities. Recall that $\mathbb{P}[|X-\mu|\geq c]\leq\frac{\mathbb{V}[X]}{c^2}$ for $c>0$.

Further recall that since $S_n\sim\text{Binomial}(n,p)$ we know $\mathbb{E}[S_n]=np$ and $\mathbb{V}[X]=np(1-p)$.

Therefore:

$$\mathbb{P}[\frac{S_n}{n}\geq p+\epsilon]=\mathbb{P}[S_n\geq n(p+\epsilon)]=\mathbb{P}[S_n-np\geq n\epsilon]\leq\mathbb{P}[S_n-np\geq n\epsilon]+\mathbb{P}[S_n-np\leq -n\epsilon]$$

$$=\mathbb{P}[|S_n-np|\geq n\epsilon]\leq\frac{np(1-p)}{n^2\epsilon^2}=\frac{p(1-p)}{n\epsilon^2}$$

If we let $n=100$, $p_i=\frac{i}{10}$, and $\epsilon=\frac{1}{10}$ this reduces to:

$$\mathbb{P}[\frac{S_n}{n}\geq p_i+\epsilon]\leq\frac{p_i(1-p_i)}{n\epsilon^2}=\frac{(i/10)(1-(i/10))}{100(1/10)^2}=\frac{1}{100}(i(10-i))$$

From the code shown **here** we got the probabilities bounds below:

i	Chebyschev bound: $(i(10-i))/100$
1	0.09
2	0.16
3	0.21
4	0.24
5	0.25
6	0.24
7	0.21
8	0.16
9	0.09

Note however that for $i=5$ we can see $\mathbb{P}[S_n-np\geq c]=\mathbb{P}[S_n-np\leq -c]$ since the distribution is symmetric about $\mu=np_i=100(5/10)=50$.

Therefore for $i=5$ which has $p_5=1/2$ we know:

$$\mathbb{P}[\frac{S_n}{n}\geq p_5+\epsilon]=\mathbb{P}[S_n\geq n(p_5+\epsilon)]=\mathbb{P}[S_n-np_5\geq n\epsilon]=\frac{1}{2}\Big(\mathbb{P}[S_n-np_5\geq n\epsilon]+\mathbb{P}[S_n-np_5\leq -n\epsilon]\Big)$$

$$=\frac{1}{2}\mathbb{P}[|S_n-np_5|\geq n\epsilon]\leq\frac{np_5(1-p_5)}{2n^2\epsilon^2}=\frac{p_5(1-p_5)}{2n\epsilon^2}$$

Showing that for $i=5$ we can cut this bound in half, this will not work for any other i because it is not true that:

$$\mathbb{P}[S_n-np_i\geq n\epsilon]=\frac{1}{2}\Big(\mathbb{P}[S_n-np_i\geq n\epsilon]+\mathbb{P}[S_n-np_i\leq -n\epsilon]\Big)$$

d.

Now we will compute the Hoeffding upper bound for these probabilities. Recall that for a random variable $Y_n = X_1 + ... + X_n$ where $a_i \leq X_i \leq b_i$ are independent we know $\mathbb{P}[Y_n - \mathbb{E}[Y_n] \geq t] \leq \exp\Big(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big)$ for $t > 0$. Further recall that since $S_n \sim \text{Binomial}(n, p)$ we know $S_n = X_1 + ... + X_n$ where $0 \leq X_i \leq 1$ are independent and $\mathbb{E}[S_n] = np$.

Therefore:

$$\mathbb{P}[\frac{S_n}{n} \geq p + \epsilon] = \mathbb{P}[S_n \geq n(p+\epsilon)] = \mathbb{P}[S_n - np \geq n\epsilon] \leq \exp\Big(-\frac{2(n\epsilon)^2}{\sum_{i=1}^n (1-0)^2}\Big) = \exp\Big(-2n\epsilon^2\Big)$$

If we let $n = 100$, $p_i = \frac{i}{10}$, and $\epsilon = \frac{1}{10}$ this reduces to:

$$\mathbb{P}[\frac{S_n}{n} \geq p_i + \epsilon] \leq \exp\Big(-2n\epsilon^2\Big) = \exp\Big(-2(100)(\frac{1}{10})^2\Big) = \exp\Big(-2\Big)$$

From the code shown **here** we got the probabilities bounds below:

i	Chebyshev bound: $\exp(-2)$
1	0.135
2	0.135
3	0.135
4	0.135
5	0.135
6	0.135
7	0.135
8	0.135
9	0.135

e.

Now we will compute the Chernoff upper bound for these probabilities. Recall that for a random variable X we know for all $a > 0$ that $\mathbb{P}[X \geq a] \leq e^{-ta}\mathbb{E}[e^{tX}]$ for $t > 0$. Further recall that $\mathbb{P}[S_n = k] = \binom{n}{k}p^k(1-p)^{n-k}$ for $k \in \{0, 1, 2, ..., n\}$.

This implies that $\mathbb{E}[e^{tX}] = \sum_{k=0}^n e^{tk}\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k}(pe^t)^k(1-p)^{n-k} = (pe^t + 1 - p)^n$

Where the last equality is a result of the binomial theorem for $(a+b)^n$ recognizing $a = pe^t$ and $b = 1-p$.

Therefore for any $t > 0$:

$$\mathbb{P}[\frac{S_n}{n} \geq p + \epsilon] = \mathbb{P}[S_n \geq n(p+\epsilon)] \leq e^{-tn(p+\epsilon)}(pe^t + 1 - p)^n$$

Now we want to find the $t > 0$ that minimizes this expression.

If $p + \epsilon = 1$ then $e^{-tn(p+\epsilon)}(pe^t + 1 - p)^n = e^{-tn}(pe^t + 1 - p)^n = (p + (1-p)e^{-t})^n \leq (p + (1-p))^n = 1$ is our bound.

Continued on next page.

If we assume $p+\epsilon<1$ (which will be apparent why we need to later) we can do this by taking the derivative with respect to t:

$$\frac{\partial}{\partial t}e^{-tn(p+\epsilon)}(pe^t+1-p)^n = -n(p+\epsilon)e^{-tn(p+\epsilon)}(pe^t+1-p)^n + ne^{-tn(p+\epsilon)}(pe^t+1-p)^{n-1}pe^t$$

Setting this equal to 0 to find critical points we get:

$$0 = -n(p+\epsilon)e^{-tn(p+\epsilon)}(pe^t+1-p)^n + ne^{-tn(p+\epsilon)}(pe^t+1-p)^{n-1}pe^t$$

$$0 = -(p+\epsilon)(pe^t+1-p)+pe^t = pe^t(1-(p+\epsilon))-(p+\epsilon)(1-p)$$

$$pe^t(1-(p+\epsilon)) = (p+\epsilon)(1-p) \qquad \Longrightarrow \qquad e^t = \frac{(p+\epsilon)(1-p)}{p(1-(p+\epsilon))}$$

$$t = \ln\Big(\frac{(p+\epsilon)(1-p)}{p(1-(p+\epsilon))}\Big) = \ln\Big(\frac{(p+\epsilon)(1-p)}{p(1-(p+\epsilon))}\Big)$$

This is in fact a minimum because the function is concave up as shown below:

$$\frac{\partial^2}{\partial t^2}e^{-tn(p+\epsilon)}(pe^t+1-p)^n = \sum_{k=0}^{n}\frac{\partial^2}{\partial t^2}\binom{n}{k}e^{-tn(p+\epsilon)}p^k e^{kt}(1-p)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}(1-p)^{n-k}p^k\frac{\partial^2}{\partial t^2}e^{t(k-n(p+\epsilon))}$$

$$= \sum_{k=0}^{n}\binom{n}{k}(1-p)^{n-k}p^k\frac{\partial}{\partial t}(k-n(p+\epsilon))e^{t(k-n(p+\epsilon))} = \sum_{k=0}^{n}\binom{n}{k}(1-p)^{n-k}p^k(k-n(p+\epsilon))^2e^{t(k-n(p+\epsilon))} > 0$$

If we let $n=100$, $p_i=\frac{i}{10}$ (for $i<9$), and $\epsilon=\frac{1}{10}$ this reduces to:

$$t = \ln\Big(\frac{(p_i+\epsilon)(1-p_i)}{p_i(1-(p_i+\epsilon))}\Big) = \ln\Big(\frac{((i/10)+(1/10))(1-(i/10))}{(i/10)(1-((i/10)+(1/10)))}\Big) = \ln\Big(\frac{(i+1)(10-i)}{i(10-i-1)}\Big)$$

Again with our bound for $i=9$ being just 1.

From the code shown **here** we got the probabilities bounds below:

i	Chernoff bound: $\ln((i+1)(10-i)/(i(10-i-1)))$ for $i<9$
1	0.811
2	0.539
3	0.442
4	0.405
5	0.405
6	0.442
7	0.539
8	0.811
9	1

2.

Recall that for a sequence of iid random variables $Y_1, Y_2, Y_3, ...$ we know by the law of large numbers that:

$$\lim_{n\to\infty} \frac{1}{n}(Y_1 + Y_2 + ... + Y_n) = \mathbb{E}[Y_1]$$

Let $X_1, X_2, ...$ be iid random variables with mean μ and variance σ^2.

Also let X belong to the same distribution.

Then we know:

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}} \sum_{i,j:1\leq i<j\leq n} (X_i - X_j)^2 = \lim_{n\to\infty} \frac{2!(n-2)!}{n!} \left(\frac{1}{2} \sum_{1\leq i,j\leq n} (X_i - X_j)^2 \right)$$

Because $(X_i - X_j)^2$ is symmetric and $(X_k - X_k)^2 = 0$. Then:

$$\lim_{n\to\infty} \frac{2!(n-2)!}{n!} \left(\frac{1}{2} \sum_{1\leq i,j\leq n} (X_i - X_j)^2 \right) = \lim_{n\to\infty} \frac{1}{n(n-1)} \sum_{1\leq i,j\leq n} (X_i - X_j)^2$$

$$= \lim_{n\to\infty} \frac{1}{n(n-1)} \sum_{1\leq i,j\leq n} (X_i - \mu + \mu - X_j)^2 = \lim_{n\to\infty} \frac{1}{n(n-1)} \sum_{1\leq i,j\leq n} (X_i-\mu)^2 + 2(X_i - \mu)(\mu - X_j) + (X_j - \mu)^2$$

$$= \lim_{n\to\infty} \frac{1}{n(n-1)} \left(\Big(\sum_{1\leq i,j\leq n} (X_i-\mu)^2 \Big) - 2\Big(\sum_{1\leq i,j\leq n} (X_i-\mu)(X_j-\mu) \Big) + \Big(\sum_{1\leq i,j\leq n} (X_j-\mu)^2 \Big) \right)$$

$$= \lim_{n\to\infty} \frac{1}{n(n-1)} \left(\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i-\mu)^2 \Big) - 2\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i-\mu)(X_j-\mu) \Big) + \Big(\sum_{i=1}^{n}\sum_{j=1}^{n} (X_j-\mu)^2 \Big) \right)$$

$$= \lim_{n\to\infty} \frac{1}{n(n-1)} \left(\Big(n\sum_{i=1}^{n} (X_i-\mu)^2 \Big) - 2\Big(\sum_{i=1}^{n} (X_i-\mu) \sum_{j=1}^{n} (X_j-\mu) \Big) + \Big(n\sum_{j=1}^{n} (X_j-\mu)^2 \Big) \right)$$

$$= \lim_{n\to\infty} 2\left(\frac{n}{n-1} \Big(\frac{1}{n} \sum_{i=1}^{n} (X_i-\mu)^2 \Big) - \frac{1}{n(n-1)} \Big(\sum_{i=1}^{n} (X_i-\mu) \Big) \Big(\sum_{j=1}^{n} (X_j-\mu) \Big) \right)$$

$$= \lim_{n\to\infty} 2\left(\frac{n}{n-1} \Big(\frac{1}{n} \sum_{i=1}^{n} (X_i-\mu)^2 \Big) - \frac{1}{n(n-1)} \Big(\sum_{i=1}^{n} (X_i-\mu) \Big)^2 \right)$$

$$= \lim_{n\to\infty} 2\left(\frac{n}{n-1} \Big(\frac{1}{n} \sum_{i=1}^{n} (X_i-\mu)^2 \Big) - \frac{n}{n^2(n-1)} \Big(\sum_{i=1}^{n} (X_i-\mu) \Big)^2 \right)$$

$$= \lim_{n\to\infty} 2\left(\frac{n}{n-1} \Big(\frac{1}{n} \sum_{i=1}^{n} (X_i-\mu)^2 \Big) - \frac{n}{n-1} \Big(\frac{1}{n} \sum_{i=1}^{n} (X_i-\mu) \Big)^2 \right)$$

Continued on next page.

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}} \sum_{i,j:1\leq i<j\leq n} (X_i - X_j)^2 = \lim_{n\to\infty} 2\left(\frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) - \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right)^2 \right)$$

$$= \lim_{n\to\infty}\left(2\frac{n}{n-1}\right)\left(\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) - \left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right)^2\right)$$

$$= \left(\lim_{n\to\infty} 2\frac{n}{n-1}\right)\left(\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) - \lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right)^2\right)$$

Since for convergent sequences (a_n) and (b_n) we know:

$$\lim_{n\to\infty} a_n b_n = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n)$$

and

$$\lim_{n\to\infty} a_n + b_n = (\lim_{n\to\infty} a_n) + (\lim_{n\to\infty} b_n)$$

and

$$\lim_{n\to\infty} ca_n = c(\lim_{n\to\infty} a_n)$$

Then:

$$\left(\lim_{n\to\infty} 2\frac{n}{n-1}\right)\left(\lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) - \lim_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right)^2\right)$$

$$= 2\left(\mathbb{E}[(X-\mu)^2] - \left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right)^2\right)$$

Where the expectation comes from the law of large numbers. Moving the limit inside of the square is allowed because $f(x) = x^2$ is continuous and for a continuous function $g(x)$ and a convergent sequence a_n we know:

$$\lim_{n\to\infty} g(a_n) = g(\lim_{n\to\infty} a_n)$$

Finally we have:

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}} \sum_{i,j:1\leq i<j\leq n} (X_i - X_j)^2 = 2\left(\mathbb{E}[(X-\mu)^2] - \left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\right)^2\right)$$

$$= 2\left(\sigma^2 - \Big(\mathbb{E}[X-\mu]\Big)^2\right) = 2\left(\sigma^2 - \Big(\mu-\mu\Big)^2\right) = 2\sigma^2 \quad \square$$

3.

We are given that $X_1, X_2, X_3, ...$ are all iid with mean $\mu = 0$ and variance σ^2. Let $S_n = X_1 + X_2 + ... + X_n$.

Firstly note that since all of the X_i's are independent we know:

$$\mathbb{V}[S_n] = \mathbb{V}[X_1 + X_2 + ... + X_n] = \mathbb{V}[X_1] + \mathbb{V}[X_2] + ... + \mathbb{V}[X_n] = n\mathbb{V}[X_1] = n\sigma^2$$

Recall Chebyschev's inequality: $\mathbb{P}[|X - \mu| \geq c] \leq \frac{\mathbb{V}[X]}{c^2}$ for $c > 0$.

a.

We can compute the first limit directly by bounding it and using the squeeze theorem:

$$0 \leq \mathbb{P}[S_n \geq 0.01n] \leq \mathbb{P}[S_n \geq 0.01n] + \mathbb{P}[S_n \leq -0.01n] = \mathbb{P}[|S_n| \geq 0.01n] = \mathbb{P}[|S_n - \mu| \geq 0.01n] \leq \frac{\mathbb{V}[S_n]}{(0.01n)^2}$$

$$= \frac{n\sigma^2}{(0.01)^2 n^2} = \frac{(100\sigma)^2}{n}$$

Therefore we know:

$$0 = \lim_{n\to\infty} 0 \leq \lim_{n\to\infty} \mathbb{P}[S_n \geq 0.01n] \leq \lim_{n\to\infty} \frac{(100\sigma)^2}{n} = 0$$

Showing by the squeeze theorem that:

$$\lim_{n\to\infty} \mathbb{P}[S_n \geq 0.01n] = 0$$

b.

By the central limit theorem we know $-\frac{S_n}{n}$ is asymptotically normal with mean $\mu = 0$ and variance $\frac{\sigma^2}{n}$.

Formally this means:

$$\lim_{n\to\infty} \mathbb{P}[\frac{-S_n}{\sigma\sqrt{n}} \leq x] = \lim_{n\to\infty} \mathbb{P}[\frac{\sqrt{n}}{\sigma}(-\frac{S_n}{n} - 0) \leq x] = \lim_{n\to\infty} \mathbb{P}[\frac{\sqrt{n}}{\sigma}((-\frac{S_n}{n}) - \mu) \leq x] = \Phi(x)$$

Where $\Phi(x)$ is the CDF of the standard normal distribution, therefore:

$$\lim_{n\to\infty} \mathbb{P}[S_n \geq 0] = \lim_{n\to\infty} \mathbb{P}[-S_n \leq 0] = \lim_{n\to\infty} \mathbb{P}[\frac{-S_n}{\sigma\sqrt{n}} \leq 0] = \Phi(0) = \frac{1}{2}$$

By the symmetry of the normal distribution.

c.

First note the following:

$$\mathbb{P}[S_n < -0.01n] \leq \mathbb{P}[S_n \leq -0.01n]$$

$$-\mathbb{P}[S_n < -0.01n] \geq -\mathbb{P}[S_n \leq -0.01n]$$

$$1 - \mathbb{P}[S_n < -0.01n] \geq 1 - \mathbb{P}[S_n \leq -0.01n]$$

From part a we know:

$$\lim_{n\to\infty} \mathbb{P}[S_n \geq 0.01n] = 0$$

Therefore:

$$1 \geq \lim_{n\to\infty} \mathbb{P}[S_n \geq -0.01n] = \lim_{n\to\infty} 1 - \mathbb{P}[S_n < -0.01n] \geq \lim_{n\to\infty} 1 - \mathbb{P}[S_n \leq -0.01n] = 1 - \lim_{n\to\infty} \mathbb{P}[S_n \leq -0.01n] = 1$$

Showing by the squeeze theorem that:

$$\lim_{n\to\infty} \mathbb{P}[S_n \geq -0.01n] = 1$$

4.

We are given the Laplace distribution has density $f_Z(z) = \frac{\lambda}{2}e^{-\lambda|z|}$ and MGF $M_Z(t) = \frac{\lambda^2}{\lambda^2-t^2}$, $\lambda > 0$.

Let $X, Y \overset{\text{iid}}{\sim} Exp(\lambda)$, then we are considering $Z - X - Y$.

Recall that $f_X(x) = \lambda e^{-\lambda x}$ for $x \geq 0$.

a.

First we will use moment generating functions. Recall that a distribution is entirely determined based on its moments and hence is entirely determined based on its moment generating function (should it exist).

Formally this means if $M_A(t) = M_B(t)$ then A and B follow the same distribution.

Computing the MGF of $Z = X - Y$ directly we have:

$$M_Z(t) = \mathbb{E}[e^tZ] = \mathbb{E}[e^{t(X-Y)}] = \mathbb{E}[e^{tX}e^{-tY}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{-tY}]$$

Where the last equality holds from the fact that X and Y are independent and so e^{tX} is independent of e^{-tY}

Now:

$$\mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}\, dx = \lambda \int_0^{\infty} e^{x(t-\lambda)}\, dx = \lambda\Big(\frac{1}{t-\lambda}e^{x(t-\lambda)}\Big|_0^{\infty}\Big)$$

$$= \frac{\lambda}{t-\lambda}\Big(\lim_{x\to\infty} e^{x(t-\lambda)} - 1\Big) = \frac{\lambda}{t-\lambda}\Big(0-1\Big) = \frac{\lambda}{\lambda - t}$$

For $t - \lambda < 0$ or equivalently $t < \lambda$ (which works fine here since MGFs consider t around a neighborhood of 0).

Similarly:

$$\mathbb{E}[e^{-tY}] = \int_{-\infty}^{\infty} e^{-ty} f_Y(y)\, dy = \int_0^{\infty} e^{-ty}\lambda e^{-\lambda y}\, dy = \lambda \int_0^{\infty} e^{-y(t+\lambda)}\, dy = \lambda\Big(\frac{-1}{t+\lambda}e^{-y(t+\lambda)}\Big|_0^{\infty}\Big)$$

$$= -\frac{\lambda}{t+\lambda}\Big(\lim_{y\to\infty} e^{-y(t+\lambda)} - 1\Big) = -\frac{\lambda}{t+\lambda}\Big(0-1\Big) = \frac{\lambda}{\lambda + t}$$

For $t + \lambda > 0$ or equivalently $t > -\lambda$ (which works fine here since MGFs consider t around a neighborhood of 0). Finally we have:

$$M_Z(t) = \mathbb{E}[e^{tX}]\mathbb{E}[e^{-tY}] = \Big(\frac{\lambda}{\lambda - t}\Big)\Big(\frac{\lambda}{\lambda + t}\Big) = \frac{\lambda^2}{\lambda^2 - t^2}$$

Which we recognize as the MGF given for the Laplace distribution.

Therefore if $X, Y \overset{\text{iid}}{\sim} Exp(\lambda)$, then $Z - X - Y$ follows the Laplace distribution $\square$

b.

Recall that the CDF of an exponential random variable with parameter μ is $F_T(t) = 1 - e^{-\mu t}$

First we will need to find the density of $-Y$, we can do this with the CDF of Y:

$\mathbb{P}[-Y \leq y] = \mathbb{P}[Y \geq -y] = \mathbb{P}[Y > -y] = 1 - \mathbb{P}[Y \leq -y] = 1 - (1 - e^{-\lambda(-y)}) = e^{\lambda y}$

For $y \leq 0$ (otherwise the probability would just be 1), therefore:

$f_{-Y}(y) = \frac{d}{dy} e^{\lambda y} = \lambda e^{\lambda y}$

For $y \leq 0$ (a rather intuitive result, we are just mirroring the function's domain).

Now we will use the convolution formula given below for $C = A + B$:

$$f_C(c) = \int_{-\infty}^{\infty} f_{A,B}(a, c - a)\, da$$

We know that X and Y in our problem are independent so $f_{X,Y}(x, y) = f_X(x) f_Y(y)$, therefore if $Z = X - Y$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)\, dx = \int_{-\infty}^{\infty} f_X(x) f_{-Y}(z - x)\, dx = \int_{0}^{\infty} \lambda e^{-\lambda x} f_{-Y}(z - x)\, dx$$

- If $z > 0$ (i.e. $z - x \leq 0$ if and only if $x \geq z > 0$):

$$f_Z(z) = \int_{0}^{\infty} \lambda e^{-\lambda x} f_{-Y}(z - x)\, dx = \int_{z}^{\infty} \lambda e^{-\lambda x} \lambda e^{\lambda(z-x)}\, dx = \lambda^2 \int_{z}^{\infty} e^{\lambda(-x+z-x)}\, dx = \lambda^2 e^{\lambda z} \int_{z}^{\infty} e^{-2\lambda x}\, dx$$
$$= \lambda^2 e^{\lambda z} \Big(\frac{-1}{2\lambda} e^{-2\lambda x} \Big|_{z}^{\infty} \Big) = \lambda^2 e^{\lambda z} \Big(\lim_{x \to \infty} \frac{-1}{2\lambda} e^{-2\lambda x} + \frac{1}{2\lambda} e^{-2\lambda z} \Big) = \lambda^2 e^{\lambda z} \Big(0 + \frac{1}{2\lambda} e^{-2\lambda z} \Big) = \frac{\lambda}{2} e^{-\lambda z}$$

- If $z \leq 0$ (i.e. $z - x < 0$ for all $x > 0$):

$$f_Z(z) = \int_{0}^{\infty} \lambda e^{-\lambda x} f_{-Y}(z - x)\, dx = \int_{0}^{\infty} \lambda e^{-\lambda x} \lambda e^{\lambda(z-x)}\, dx = \lambda^2 \int_{0}^{\infty} e^{\lambda(-x+z-x)}\, dx = \lambda^2 e^{\lambda z} \int_{0}^{\infty} e^{-2\lambda x}\, dx$$
$$= \lambda^2 e^{\lambda z} \Big(\frac{-1}{2\lambda} e^{-2\lambda x} \Big|_{0}^{\infty} \Big) = \lambda^2 e^{\lambda z} \Big(\lim_{x \to \infty} \frac{-1}{2\lambda} e^{-2\lambda x} + \frac{1}{2\lambda} \Big) = \lambda^2 e^{\lambda z} \Big(0 + \frac{1}{2\lambda} \Big) = \frac{\lambda}{2} e^{\lambda z}$$

Therefore:

$$f_Z(z) = \begin{cases} \frac{\lambda}{2} e^{\lambda z} & \text{for} \quad z \leq 0 \\ \frac{\lambda}{2} e^{-\lambda z} & \text{for} \quad z > 0 \end{cases} = \frac{\lambda}{2} e^{-\lambda |z|} \text{ for all } z \in \mathbb{R}$$

Which we recognize as the density given for the Laplace distribution.

Therefore if $X, Y \overset{\text{iid}}{\sim} Exp(\lambda)$, then $Z - X - Y$ follows the Laplace distribution $\square$

5.

We are given the following PDF for X:

$$f_X(x) = \begin{cases} \frac{2}{9} & \text{for } 0 \le x \le 1 \\ \frac{4-|4-2x|}{9} & \text{for } 1 < x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

a.

First we will show that this is indeed a PDF.

Note that $|4-2x| \le 4$ for $1 < x \le 4$, therefore $4 - |4-2x| \ge 0$ for $1 < x \le 4$. Clearly $2/9 > 0$.

So we can clearly see that $f_X(x) \ge 0$ for all $x \in \mathbb{R}$.

Now see the following:

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 \frac{2}{9}\,dx + \int_1^2 \frac{4-|4-2x|}{9}\,dx + \int_2^4 \frac{4-|4-2x|}{9}\,dx$$

$$= \frac{2}{9} + \int_1^2 \frac{4-(4-2x)}{9}\,dx + \int_2^4 \frac{4+(4-2x)}{9}\,dx = \frac{2}{9} + \frac{2}{9}\int_1^2 x\,dx + \frac{2}{9}\int_2^4 4-x\,dx$$

$$= \frac{2}{9}\left(1 + \left(\frac{x^2}{2}\Big|_1^2\right) + \left(4x - \frac{x^2}{2}\Big|_2^4\right)\right) = \frac{2}{9}\left(1 + \left(2 - \frac{1}{2}\right) + \left(16 - 8 - 8 + 2\right)\right)$$

$$= \frac{2}{9}\left(1 + \frac{3}{2} + 2\right) = \left(\frac{2}{9}\right)\left(\frac{9}{2}\right) = 1$$

Therefore $f_X(x)$ is a PDF since it is non-negative and integrates to 1 $\square$

Part b on next page.

b.

Now we will find the MGF (for $t \neq 0$):

$$M_X(t) = \mathbb{E}[e^{Xt}] = \int_{-\infty}^{\infty} e^{xt} f_X(x)\, dx = \int_0^1 \frac{2}{9} e^{xt}\, dx + \int_1^2 \frac{4 - |4-2x|}{9} e^{xt}\, dx + \int_2^4 \frac{4-|4-2x|}{9} e^{xt}\, dx$$

$$= \frac{2}{9}\Big(\frac{1}{t} e^{xt}\Big|_0^1\Big) + \int_1^2 \frac{4-(4-2x)}{9} e^{xt}\, dx + \int_2^4 \frac{4+(4-2x)}{9} e^{xt}\, dx$$

$$= \frac{2(e^t - 1)}{9t} + \frac{2}{9}\int_1^2 x e^{xt}\, dx + \frac{2}{9}\int_2^4 (4-x) e^{xt}\, dx$$

$$= \frac{2(e^t-1)}{9t} + \frac{2}{9}\int_2^4 4e^{xt}\, dx + \frac{2}{9}\int_1^2 xe^{xt}\, dx - \frac{2}{9}\int_2^4 xe^{xt}\, dx$$

$$= \frac{2(e^t-1)}{9t} + \frac{2}{9}\Big(\frac{4}{t} e^{xt}\Big|_2^4\Big) + \frac{2}{9}\int_1^2 xe^{xt}\, dx - \frac{2}{9}\int_2^4 xe^{xt}\, dx$$

$$= \frac{2(e^t-1)}{9t} + \frac{8(e^{4t} - e^{2t})}{9t} + \frac{2}{9}\int_1^2 xe^{xt}\, dx - \frac{2}{9}\int_2^4 xe^{xt}\, dx$$

Now we will use substitution to solve the remaining integrals:

Let $u = x$ and $\frac{dv}{dx} = e^{xt}$ then $\frac{du}{dx} = 1$ and $v = \frac{1}{t}e^{xt}$, then:

$$\int_a^b xe^{xt}\, dx = \int_a^b u\frac{dv}{dx}\, dx = uv\Big|_a^b - \int_a^b v\frac{du}{dx}\, dx = \frac{x}{t}e^{xt}\Big|_a^b - \int_a^b \frac{1}{t}e^{xt}\, dx = \frac{be^{bt} - ae^{at}}{t} - \Big(\frac{1}{t^2}e^{xt}\Big|_a^b\Big)$$

$$= \frac{be^{bt} - ae^{at}}{t} - \frac{e^{bt} - e^{at}}{t^2}$$

Therefore for $t \neq 0$ we have:

$$M_X(t) = \frac{2(e^t-1)}{9t} + \frac{8(e^{4t}-e^{2t})}{9t} + \frac{2}{9}\int_1^2 xe^{xt}\, dx - \frac{2}{9}\int_2^4 xe^{xt}\, dx$$

$$= \frac{2}{9}\left(\frac{e^t - 1}{t} + \frac{4(e^{4t}-e^{2t})}{t} + \frac{2e^{2t} - e^t}{t} - \frac{e^{2t}-e^t}{t^2} - \frac{4e^{4t} - 2e^{2t}}{t} + \frac{e^{4t} - e^{2t}}{t^2}\right)$$

$$= \frac{2}{9}\left(\frac{e^t - 1 + 4e^{4t} - 4e^{2t} + 2e^{2t} - e^t - 4e^{4t} + 2e^{2t}}{t} - \frac{e^{2t} - e^t - e^{4t} + e^{2t}}{t^2}\right)$$

$$= \frac{2}{9}\left(\frac{e^{4t} + e^t - 2e^{2t} - t}{t^2}\right)$$

If $t = 0$ then $M_X(t) = \mathbb{E}[e^{Xt}] = \mathbb{E}[e^{X0}] = \mathbb{E}[1] = 1$. Therefore the MGF for X if given by:

$$M_X(t) = \mathbb{E}[e^{Xt}] = \begin{cases} \frac{2}{9}\left(\frac{e^{4t}+e^t-2e^{2t}-t}{t^2}\right) & \text{for } \ t \neq 0 \\ 1 & \text{for } \ t = 0 \end{cases} \qquad \square$$

c.

Recall that

$$M_X(t) = \mathbb{E}[e^{Xt}] = \begin{cases} \frac{2}{9}\left(\frac{e^{4t}+e^t-2e^{2t}-t}{t^2}\right) & \text{for } \ t \neq 0 \\ 1 & \text{for } \ t = 0 \end{cases} \qquad \square$$

First note that $M_X(t)$ satisfies the $\frac{0}{0}$ condition for L'hopital's rule twice:

$$\lim_{t\to 0} \frac{2}{9}\left(\frac{e^{4t}+e^t-2e^{2t}-t}{t^2}\right) = \frac{2}{9}\lim_{t\to 0}\frac{\frac{d}{dt}(e^{4t}+e^t-2e^{2t}-t)}{\frac{d}{dt}t^2} = \frac{2}{9}\lim_{t\to 0}\frac{4e^{4t}+e^t-4e^{2t}-1}{2t}$$

$$= \frac{2}{9}\lim_{t\to 0}\frac{\frac{d}{dt}(4e^{4t}+e^t-4e^{2t}-1)}{\frac{d}{dt}2t} = \frac{2}{9}\lim_{t\to 0}\frac{16e^{4t}+e^t-8e^{2t}}{2} = \left(\frac{2}{9}\right)\left(\frac{16+1-8}{2}\right) = \left(\frac{2}{9}\right)\left(\frac{9}{2}\right) = 1$$

Therefore $M_X(t)$ is continuous for all $t \in \mathbb{R}$.

Because $M_X(t)$ is continuous at $t=0$ we know for all $t \in \mathbb{R}$ that:

$$M_X(t) = \frac{2}{9t^2}\left(e^{4t}+e^t-2e^{2t}-t\right) = \frac{2}{9t^2}\left(-t+\sum_{n=0}^{\infty}\frac{(4t)^n}{n!}+\sum_{n=0}^{\infty}\frac{t^n}{n!}-2\sum_{n=0}^{\infty}\frac{(2t)^n}{n!}\right)$$

$$= \frac{2}{9t^2}\left(-t+\sum_{n=0}^{\infty}\frac{4^nt^n+t^n-2^{n+1}t^n}{n!}\right) = \frac{2}{9t^2}\left(-t+\sum_{n=0}^{\infty}t^n\frac{4^n+1-2^{n+1}}{n!}\right)$$

$$= \frac{2}{9t^2}\left(-t+t+\sum_{n=2}^{\infty}t^n\frac{4^n+1-2^{n+1}}{n!}\right) = \frac{2}{9t^2}\sum_{n=2}^{\infty}t^n\frac{4^n+1-2^{n+1}}{n!}$$

$$= \frac{2}{9}\sum_{n=2}^{\infty}t^{n-2}\frac{4^n+1-2^{n+1}}{n!}$$

Clearly $M_X(t) < \infty$ for all $t \in (-\epsilon, \epsilon)$ for some $\epsilon > 0$ (in fact for all $\epsilon > 0$ in this case).

Therefore we know $\frac{d^k}{dt^k}M_X(t)\Big|_{t=0} = \mathbb{E}[X^k]$ for all $k \in \mathbb{N}$.

First computing $\mathbb{E}[X]$:

$$\mathbb{E}[X] = \frac{d}{dt}M_X(t)\Big|_{t=0} = \frac{d}{dt}\frac{2}{9}\sum_{n=2}^{\infty}t^{n-2}\frac{4^n+1-2^{n+1}}{n!}\Bigg|_{t=0}$$

$$= \frac{2}{9}\sum_{n=2}^{\infty}\frac{d}{dt}t^{n-2}\frac{4^n+1-2^{n+1}}{n!}\Bigg|_{t=0} = \frac{2}{9}\sum_{n=3}^{\infty}(n-2)t^{n-3}\frac{4^n+1-2^{n+1}}{n!}\Bigg|_{t=0}$$

$$= \left(\frac{2}{9}\right)\left(\frac{4^3+1-2^4}{3!}\right) = \left(\frac{2}{9}\right)\left(\frac{64+1-16}{6}\right) = \left(\frac{2}{9}\right)\left(\frac{49}{6}\right) = \frac{49}{27}$$

Continued on next page.

Now computing $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = \frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = \frac{d^2}{dt^2} \frac{2}{9} \sum_{n=2}^{\infty} t^{n-2} \frac{4^n + 1 - 2^{n+1}}{n!}\Bigg|_{t=0}$$

$$= \frac{2}{9} \sum_{n=2}^{\infty} \frac{d^2}{dt^2} t^{n-2} \frac{4^n + 1 - 2^{n+1}}{n!}\Bigg|_{t=0} = \frac{2}{9} \sum_{n=3}^{\infty} (n-2) \frac{d}{dt} t^{n-3} \frac{4^n + 1 - 2^{n+1}}{n!}\Bigg|_{t=0}$$

$$= \frac{2}{9} \sum_{n=4}^{\infty} (n-2)(n-3) t^{n-4} \frac{4^n + 1 - 2^{n+1}}{n!}\Bigg|_{t=0} = \Big(\frac{2}{9}\Big)\Big(2\Big)\Big(\frac{4^4 + 1 - 2^5}{4!}\Big)$$

$$= \Big(\frac{4}{9}\Big)\Big(\frac{256 + 1 - 32}{24}\Big) = \Big(\frac{4}{9}\Big)\Big(\frac{225}{24}\Big) = \frac{25}{6}$$

Therefore we know:

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{25}{6} - \frac{49^2}{27^2} = \frac{(25)(27^2) - (6)(49^2)}{(6)(27^2)} = \frac{18225 - 14406}{4374} = \frac{1273}{1458}$$

Giving us our final answer:

$$\mathbb{E}[X] = \frac{49}{27} \qquad\qquad \mathbb{V}[X] = \frac{1273}{1458}$$

d.

Again we will use the fact that $\frac{d^k}{dt^k}M_X(t)\Big|_{t=0} = \mathbb{E}[X^k]$ for all $k \in \mathbb{N}$.

Recall:

$$M_X(t) = \frac{2}{9}\sum_{n=2}^{\infty} t^{n-2}\frac{4^n + 1 - 2^{n+1}}{n!}$$

For $m < k$ when we take the kth derivative of at^m for $a \in \mathbb{R}$ we get:

$$\frac{d^k}{dt^k}at^m = a\frac{d^{(k-m)}}{dt^{(k-m)}}m!t^{(m-m)} = a\frac{d^{(k-m)}}{dt^{(k-m)}}m! = 0$$

Since the derivative of a constant is 0.

Then if we take the kth derivative of at^k for $a \in \mathbb{R}$ we get:

$$\frac{d^k}{dt^k}at^k = a(k!t^{(k-k)}) = a(k!)$$

For $m > k$ when we take the kth derivative of at^m and evaluate at $t = 0$ for $a \in \mathbb{R}$ we get:

$$\frac{d^k}{dt^k}at^m\Big|_{t=0} = am(m-1)...(m-k+1)t^{m-k}\Big|_{t=0} = a\frac{m!}{(m-k)!}t^{m-k}\Big|_{t=0} = 0$$

When we take the kth derivative of $M_X(t)$ and evaluate at $t = 0$ we will get:

$$\frac{d^k}{dt^k}M_X(t)\Big|_{t=0} = \frac{d^k}{dt^k}\frac{2}{9}\sum_{n=2}^{\infty} t^{n-2}\frac{4^n + 1 - 2^{n+1}}{n!}\Bigg|_{t=0} = \frac{2}{9}\sum_{n=2}^{\infty}\frac{d^k}{dt^k}t^{n-2}\frac{4^n + 1 - 2^{n+1}}{n!}\Bigg|_{t=0}$$

$$= \left(\frac{2}{9}\right)k!\frac{4^{k+2} + 1 - 2^{k+3}}{(k+2)!} = \frac{2(4^{k+2} + 1 - 2^{k+3})}{9(k+2)(k+1)}$$

Since only the term where $k = n-2$ (i.e. $n = k+2$) will remain due to the results above.

Giving us the final result:

$$\mathbb{E}[X^k] = \frac{d^k}{dt^k}M_X(t)\Big|_{t=0} = \frac{2(4^{k+2} + 1 - 2^{k+3})}{9(k+2)(k+1)}$$

6.

We are letting $X_1, X_2, ..., X_n$ be independent and $S_n = X_1 + X_2 + ... + X_n$.

a.

First we are considering $X_i \sim N(\mu_i, \sigma_i^2)$.

Finding the MGF for a normal distribution we have (letting $X \sim N(\mu, \sigma^2)$):

$$M_X(t) = \mathbb{E}[e^{Xt}] = \int_{-\infty}^{\infty} e^{xt} f_X(x)\, dx = \int_{-\infty}^{\infty} \frac{e^{xt}}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{xt - \frac{x^2 - 2\mu x + \mu^2}{2\sigma^2}}\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2 - 2\mu x - 2\sigma^2 t x + \mu^2}{2\sigma^2}}\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2 - 2x(\mu+\sigma^2 t) + \mu^2}{2\sigma^2}}\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2 - 2x(\mu+\sigma^2 t) + (\mu+\sigma^2 t)^2 - (\mu+\sigma^2 t)^2 + \mu^2}{2\sigma^2}}\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2 - 2x(\mu+\sigma^2 t) + (\mu+\sigma^2 t)^2}{2\sigma^2}} e^{\frac{(\mu+\sigma^2 t)^2 - \mu^2}{2\sigma^2}}\, dx$$

$$= \frac{e^{\frac{(\mu+\sigma^2 t)^2 - \mu^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2 - 2x(\mu+\sigma^2 t) + (\mu+\sigma^2 t)^2}{2\sigma^2}}\, dx = \frac{e^{\frac{(\mu+\sigma^2 t)^2 - \mu^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}}\, dx$$

$$= e^{\frac{(\mu+\sigma^2 t)^2 - \mu^2}{2\sigma^2}} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}}\, dx = e^{\frac{(\mu+\sigma^2 t)^2 - \mu^2}{2\sigma^2}} = e^{\frac{\mu^2 + 2\mu\sigma^2 t + \sigma^4 t^2 - \mu^2}{2\sigma^2}}$$

$$= e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}} = e^{\frac{2\mu t + \sigma^2 t^2}{2}} = e^{\mu t} e^{\frac{\sigma^2 t^2}{2}}$$

Where the equality getting rid of the integrand holds by noticing that the function inside the integrand is the density of a $N(\mu + \sigma^2 t, \sigma^2)$ so it must integrate to 1. This MGF will apply to all of our X_i we simply need to replace μ with μ_i and σ^2 with σ_i^2.

Recall that the expectation of the product of independent random variables is the product of their expectations. Then since all the X_i are independent all of the $e^{X_i t}$ are independent, so we have:

$$M_{S_n}(t) = \mathbb{E}[e^{S_n t}] = \mathbb{E}[e^{(X_1 + X_2 + ... + X_n)t}] = \mathbb{E}[e^{X_1 t + X_2 t + ... + X_n t}] = \mathbb{E}[e^{X_1 t} e^{X_2 t} ... \, e^{X_n t}]$$

$$= \mathbb{E}[e^{X_1 t}]\mathbb{E}[e^{X_2 t}] ... \mathbb{E}[e^{X_n t}] = e^{\mu_1 t} e^{\frac{\sigma_1^2 t^2}{2}} e^{\mu_2 t} e^{\frac{\sigma_2^2 t^2}{2}} ... \, e^{\mu_n t} e^{\frac{\sigma_n^2 t^2}{2}} = e^{\mu_1 t} e^{\mu_2 t} ... \, e^{\mu_n t} e^{\frac{\sigma_1^2 t^2}{2}} e^{\frac{\sigma_2^2 t^2}{2}} ... \, e^{\frac{\sigma_n^2 t^2}{2}}$$

$$= e^{(\mu_1 + \mu_2 + ... + \mu_n)t} e^{\frac{(\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2)t^2}{2}}$$

Which we recognize as the MGF of a $N(\mu_1 + \mu_2 + ... + \mu_n, \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2)$

As before recall that a distribution is entirely determined from its moment generating function (should it exist).

Formally this means if $M_A(t) = M_B(t)$ then A and B follow the same distribution.

Therefore if $X_1, X_2, ..., X_n$ are independent with $X_i \sim N(\mu_i, \sigma_i^2)$ then:

$S_n = X_1 + X_2 + ... + X_n \sim N(\mu_1 + \mu_2 + ... + \mu_n, \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2)$ $\square$

b.

Now we are considering $X_i \sim \text{Gamma}(r_i, \lambda)$.

Finding the MGF for a gamma distribution we have (letting $X \sim \text{Gamma}(\alpha, \beta)$):

For $\beta - t > 0$ i.e. $t < \beta$ (which works fine here since MGFs consider t around a neighborhood of 0).

$$M_X(t) = \mathbb{E}[e^{Xt}] = \int_{-\infty}^{\infty} e^{xt} f_X(x)\, dx = \int_0^{\infty} e^{xt} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} e^{-x(\beta-t)}\, dx$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} \frac{(\beta-t)^\alpha}{(\beta-t)^\alpha} e^{-x(\beta-t)}\, dx = \frac{\beta^\alpha}{(\beta-t)^\alpha} \int_0^{\infty} \frac{(\beta-t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x(\beta-t)}\, dx$$

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

Where the equality getting rid of the integrand holds by noticing that the function inside the integrand is the density of a $\text{Gamma}(\alpha, \beta - t)$ so it must integrate to 1. This MGF will apply to all of our X_i we simply need to replace α with r_i and β with λ.

Recall that the expectation of the product of independent random variables is the product of their expectations.

Then since all the X_i are independent all of the $e^{X_i t}$ are independent, so we have:

$$M_{S_n}(t) = \mathbb{E}[e^{S_n t}] = \mathbb{E}[e^{(X_1+X_2+...+X_n)t}] = \mathbb{E}[e^{X_1 t + X_2 t + ... + X_n t}] = \mathbb{E}[e^{X_1 t} e^{X_2 t} ... \, e^{X_n t}]$$

$$= \mathbb{E}[e^{X_1 t}]\mathbb{E}[e^{X_2 t}]...\mathbb{E}[e^{X_n t}] = \left(\frac{\lambda}{\lambda-t}\right)^{r_1} \left(\frac{\lambda}{\lambda-t}\right)^{r_2} ... \left(\frac{\lambda}{\lambda-t}\right)^{r_n} = \left(\frac{\lambda}{\lambda-t}\right)^{r_1+r_2+...+r_n}$$

Which we recognize as the MGF of a $\text{Gamma}(r_1 + r_2 + ... + r_n, \lambda)$

As before recall that a distribution is entirely determined from its moment generating function (should it exist).

Formally this means if $M_A(t) = M_B(t)$ then A and B follow the same distribution.

Therefore if $X_1, X_2, ..., X_n$ are independent with $X_i \sim \text{Gamma}(r_i, \lambda)$ then:

$S_n = X_1 + X_2 + ... + X_n \sim \text{Gamma}(r_1 + r_2 + ... + r_n, \lambda)$ $\square$

c.

Now we are considering $X_i = Z_i^2$ where $Z_i \sim N(0,1)$.

Finding the MGF for $X = Z^2$ where $Z \sim N(0,1)$:

$$M_X(t) = \mathbb{E}[e^{Xt}] = \mathbb{E}[e^{Z^2 t}] = \int_{-\infty}^{\infty} e^{z^2 t} f_Z(z)\,dz = \int_{-\infty}^{\infty} \frac{e^{z^2 t}}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\,dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2(\frac{1}{2}-t)}\,dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}(1-2t)}\,dz$$

First note this integral only converges for $1-2t>0$

(i.e. $t<\frac{1}{2}$ which works fine here since MGFs consider t around a neighborhood of 0).

Now let $\sigma^2 = \frac{1}{1-2t}$ (again taking $t<\frac{1}{2}$), then:

$$M_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2\sigma^2}}\,dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\sigma}{\sigma} e^{-\frac{z^2}{2\sigma^2}}\,dz$$

$$= \sigma \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}\,dz = \sigma = \frac{1}{\sqrt{1-2t}}$$

Where the equality getting rid of the integrand holds by noticing that the function inside the integrand is the density of a $N(0,\sigma^2)$ so it must integrate to 1. This MGF will apply to all of our X_i.

Recall that the expectation of the product of independent random variables is the product of their expectations.

Then since all the X_i are independent all of the $e^{X_i t}$ are independent, so we have:

$$M_{S_n}(t) = \mathbb{E}[e^{S_n t}] = \mathbb{E}[e^{(X_1+X_2+...+X_n)t}] = \mathbb{E}[e^{X_1 t + X_2 t + ... + X_n t}] = \mathbb{E}[e^{X_1 t} e^{X_2 t} ... \, e^{X_n t}]$$

$$= \mathbb{E}[e^{X_1 t}]\mathbb{E}[e^{X_2 t}]...\mathbb{E}[e^{X_n t}] = \Big(\frac{1}{\sqrt{1-2t}}\Big)\Big(\frac{1}{\sqrt{1-2t}}\Big)...\Big(\frac{1}{\sqrt{1-2t}}\Big)$$

$$= \Big(\frac{1}{\sqrt{1-2t}}\Big)^n = \frac{1}{\sqrt{(1-2t)^n}} = \Big(\frac{1}{1-2t}\Big)^{\frac{n}{2}} = \Big(\frac{1/2}{1/2-t}\Big)^{\frac{n}{2}}$$

Which we recognize as the MGF of a Gamma$(\frac{n}{2}, \frac{1}{2})$

As before recall that a distribution is entirely determined from its moment generating function (should it exist).

Formally this means if $M_A(t) = M_B(t)$ then A and B follow the same distribution.

Therefore if $X_1, X_2, ..., X_n$ are independent with $X_i = Z_i^2$ where $Z_i \sim N(0,1)$ then:

$S_n = X_1 + X_2 + ... + X_n \sim$ Gamma$(\frac{n}{2}, \frac{1}{2})$ $\square$

This is also called the Chi-Squared distribution with n degrees of freedom so we can also write:

$S_n = X_1 + X_2 + ... + X_n \sim \chi^2_n$ $\square$

Code for Problem 1:

1.a. code

```
n <- 100
eps <- 1/10
for (i in 1:9){
  s <- 0
  p <- i/10
  k <- 10*(i+1)
  while (k <= 100){
    s <- s + choose(n, k)*(p^k)*((1-p)^(n-k))
    k <- k + 1
  }
  print(sprintf("Probability for i = %i: %s", i, signif(s, 3)))
}
```

```
## [1] "Probability for i = 1: 0.00198"
## [1] "Probability for i = 2: 0.0112"
## [1] "Probability for i = 3: 0.021"
## [1] "Probability for i = 4: 0.0271"
## [1] "Probability for i = 5: 0.0284"
## [1] "Probability for i = 6: 0.0248"
## [1] "Probability for i = 7: 0.0165"
## [1] "Probability for i = 8: 0.0057"
## [1] "Probability for i = 9: 2.66e-05"
```

1.b. code

```
for (i in 1:9){
  p_bound <- i/(i+1)
  print(sprintf("Markov probability bound for i = %i: %s", i, signif(p_bound, 3)))
}
```

```
## [1] "Markov probability bound for i = 1: 0.5"
## [1] "Markov probability bound for i = 2: 0.667"
## [1] "Markov probability bound for i = 3: 0.75"
## [1] "Markov probability bound for i = 4: 0.8"
## [1] "Markov probability bound for i = 5: 0.833"
## [1] "Markov probability bound for i = 6: 0.857"
## [1] "Markov probability bound for i = 7: 0.875"
## [1] "Markov probability bound for i = 8: 0.889"
## [1] "Markov probability bound for i = 9: 0.9"
```

1.c. code

```
for (i in 1:9){
  p_bound <- i*(10-i)/100
  print(sprintf("Markov probability bound for i = %i: %s", i, signif(p_bound, 3)))
}
```

```
## [1] "Markov probability bound for i = 1: 0.09"
## [1] "Markov probability bound for i = 2: 0.16"
## [1] "Markov probability bound for i = 3: 0.21"
## [1] "Markov probability bound for i = 4: 0.24"
## [1] "Markov probability bound for i = 5: 0.25"
## [1] "Markov probability bound for i = 6: 0.24"
## [1] "Markov probability bound for i = 7: 0.21"
## [1] "Markov probability bound for i = 8: 0.16"
## [1] "Markov probability bound for i = 9: 0.09"
```

1.d. code

```
print(sprintf("Hoeffding probability bound: %s", signif(exp(-2), 3)))
```

```
## [1] "Hoeffding probability bound: 0.135"
```

1.e. code

```
for (i in 1:9){
  p_bound <- log(((i+1)*(10-i))/(i*(10-i-1)))
  if (i == 9){
    p_bound <- 1
  }
  print(sprintf("Markov probability bound for i = %i: %s", i, signif(p_bound, 3)))
}
```

```
## [1] "Markov probability bound for i = 1: 0.811"
## [1] "Markov probability bound for i = 2: 0.539"
## [1] "Markov probability bound for i = 3: 0.442"
## [1] "Markov probability bound for i = 4: 0.405"
## [1] "Markov probability bound for i = 5: 0.405"
## [1] "Markov probability bound for i = 6: 0.442"
## [1] "Markov probability bound for i = 7: 0.539"
## [1] "Markov probability bound for i = 8: 0.811"
## [1] "Markov probability bound for i = 9: 1"
```