1. Midterm 01

2. Consider a random variable Z with transform:

$$M_Z(s) = \frac{a - 3s}{s^2 - 6s + 8}$$

(a) Find the numerical value for the parameter a.

Evaluating any MGF at 0 will yield 1 (which is a simple consequence of the Taylor Series expansion of e^x). So:

$$M_Z(0) = \frac{a}{8} = 1$$

So $a = 8$.

(b) Find $P(Z \geq 0.5)$.

First we have to find the PDF of Z. Notice that the MGF is just a linear combination of two exponential MGFs:

$$\frac{8 - 3s}{(4 - s)(2 - s)}$$

$$= \frac{2(2 - s)}{(4 - s)(2 - s)} + \frac{4 - s}{(2 - s)(4 - s)}$$

$$= \frac{1}{2}\left(\frac{4}{4 - s}\right) + \frac{1}{2}\left(\frac{2}{2 - s}\right)$$

So we see the PDF of Z is just:

$$f_Z(x) = \frac{1}{2}f_1(x) + \frac{1}{2}f_2(x)$$

$$f_Z(x) = 2e^{-4x} + e^{-2x}$$

Then, the CDF is given by:

$$F_Z(x) = \int_{x=0}^{x} 2e^{-2x} + e^{-2x} dx$$

After evaluating the integral, we get:

$$F_Z(x) = -\frac{1}{2}(e^{-4x} + e^{-2x}) + 1$$

Finally:

$$P(Z \geq 0.5) = 1 - F_Z(0.5) = 1 - \left(-\frac{1}{2}(e^{-2} + e^{-1}) + 1\right)$$

$$P(Z \geq 0.5) = \frac{1}{2}(e^{-2} + e^{-1}) \approx 25\%$$

(c) Find $E[Z]$ by using the probability distribution of Z.

$$E[Z] = \int_{x=0}^{\infty} 2xe^{-4x} + xe^{-2x} dx$$

Splitting up the two terms and integrating by parts we get:

$$E[Z] = \frac{3}{8}$$

(d) Find $E[Z]$ by using the transform of Z and without explicitly using the probability distribution of Z.

We can get this by evaluating the first derivative of our MGF at 0.

$$M'_Z(s) = \frac{3s^2 - 16s + 24}{(s^2 - 6s + 8)^2}$$

So then:

$$M'_Z(0) = \frac{24}{64} = \frac{3}{8}$$

(e) Find Var(Z) by using the probability distribution of Z.

We can use the definition that $\text{Var}(Z) = E[X^2] - E[X]^2$. So we just need to find $E[X^2]$ with our PDF. Like last time, this is just:

$$E[Z^2] = \int_{x=0}^{\infty} 2x^2 e^{-4x} + x^2 e^{-2x} dx$$

Again, with integration by parts we get:

$$E[Z^2] = \frac{5}{16}$$

Then:

$$Var(Z) = \frac{5}{16} - (\frac{3}{8})^2 = \frac{11}{64}$$

(f) Find Var(Z) by using the transform of Z and without explicitly using the probability distribution of Z.

Now, we need to evaluate the MGF's **second** derivative at 0. So we have:

$$M''_Z(s) = \frac{-2(3s^2 - 24s^2 + 72s - 80)}{(s^2 - 6s + 8)^3}$$

$$M''_Z(0) = \frac{-2(-80)}{8^3} = \frac{5}{16}$$

Then:

$$Var(Z) = \frac{5}{16} - (\frac{3}{8})^2 = \frac{11}{64}$$

3. Let X, Y, and Z be independent random variables. X is Bernoulli with $p = \frac{1}{4}$. Y is exponential with parameter 3. Z is Poisson with parameter 5.

(a) Find the transform of $5Z + 1$.

$$E[e^{(5Z+1)t}] = \Sigma_{z=0}^{\infty} e^{5zt+t} (\frac{5^z e^{-5}}{z!})$$

$$= e^{t-5} \Sigma_z \frac{(5e^{5t})^z}{z!}$$

This, luckily, is the Taylor Series expansion of e^x for $x = 5e^{5t}$. So we get:

$$M_{5Z+1}(t) = e^{t-5} e^{5e^{5t}} = e^{5e^{5t}+t-5}$$

(b) Find the transform of $X + Y$.

Since X and Y are independent, we know that the MGF of the sum is the product of the MGFs. This is:

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

$$M_{X+Y}(t) = (\frac{3}{4} + \frac{1}{4} e^t)(\frac{3}{3-t})$$

(c) Consider the new random variable $U = XY + (1 - X)Z$. Find the transform associated with U.

TODO!!!

4. In class, we learned some inequalities such as the Markov Inequality, the Chebyshev Inequality, and the Chernoff Bound. In this problem, we will derive an inequality, which is a special case of Chernoff bound, using a simple counting method.
Suppose $X_1, \cdots, X_n$ are i.i.d. Bernoulli random variables with $P(X_i = 1) = \frac{1}{2}$.

(a) First, use the Chebyshev inequality to show that for any $\epsilon > 0$,

$$P(\Sigma_{i=1}^n X_i \geq \frac{n}{2}(1 + \epsilon)) \leq \frac{1}{\epsilon^2 n}$$

Let $\hat{X} = \frac{\Sigma X_i}{n}$. Then, we know that $E[\hat{X}] = p$, and $Var(\hat{X}) = \frac{1}{4n}$, since $E[X_i] = p$ and $Var(X_i) = \frac{1}{4}$. From the Chebyshev inequality, we know that:

$$P(\left| \hat{X} - \frac{1}{2} \right| > \frac{\epsilon}{2}) \leq \frac{\frac{1}{4}(4)}{n\epsilon^2}$$

For any $\epsilon > 0$. But if we multiply the inequality inside the probability by n, we get:

$$P(\left| \Sigma X_i - n\frac{1}{2} > \frac{n\epsilon}{2} \right|) = P(\Sigma X_i > \frac{n}{2}(\epsilon + 1))$$

So putting it together we finally get:

$$P(\Sigma_{i=1}^n X_i \geq \frac{n}{2}(1 + \epsilon)) \leq \frac{1}{\epsilon^2 n}$$

(b) Let M be the event that $X_1 = X_2 = \cdots = X_m = 1, m < n$. Show that for an integer $k(m \le k \le n)$,

$$P(M|\Sigma_{i=1}^n X_i = k) \ge (\frac{k-m}{n-m})^m$$

Just from definitions we get:

$$P(M|\Sigma_{i=1}^n X_i = k) = \frac{P(M)P(\Sigma X_i = k|M)}{P(\Sigma X_i = k)}$$

$$= \frac{(p^m)(\binom{n-m}{k-m}p^{k-m}(1-p)^{n-k})}{\binom{n}{k}p^k(1-p)^{n-k}}$$

Since $p = (1-p)$, we see that all of the p and $1-p$ terms magically cancel giving us:

$$= \frac{\binom{n-m}{k-m}}{\binom{n}{k}} = \frac{k(k-1)\cdots(k-m+1)}{n(n-1)\cdots(n-m+1)}$$

From simple reasoning, we know that there are m terms in the multiplications of the numerator and denominator. But every individual term is greater than $k-m$ and $n-m$, so all m of those multiplications **must** be greater than $(\frac{k-m}{n-m})^m$. So the inequality stands:

$$P(M|\Sigma_{i=1}^n X_i = k) \ge (\frac{k-m}{n-m})^m$$

As for the second inequality, it's a natural extension of the first. If there are more ways for the sum of all X_i to be k, then there are more ways in which the first m terms are 1, so we know that $P(M|\Sigma_{i=1}^n X_i \ge k) > P(M|\Sigma_{i=1}^n X_i = k)$. Thus, $P(M|\Sigma_{i=1}^n X_i \ge k) \ge (\frac{k-m}{n-m})^m$ by transitivity.

(c) For simplicity, we assume that $\frac{\epsilon n}{4}$ is an integer and let $m = \frac{\epsilon n}{4}$. Let G be the event that $\Sigma X_i \ge \frac{n}{2}(1+\epsilon)$. Show that

$$P(M|G) \ge (\frac{1}{2} + \frac{\epsilon}{4})^m$$

From part b we know that:

$$P(M|G) \ge (\frac{\frac{n}{2}(1+\epsilon) - \frac{\epsilon n}{4}}{n - \frac{\epsilon n}{4}})^m$$

Simplifying:

$$P(M|G) \ge (\frac{2+\epsilon}{4-\epsilon})$$

We know that if we increase the denominator, we are decreasing the overall number (because we're working with all positive numbers). This applies with repeated increasing of the denominator, so we can do:

$$P(M|G) \ge (\frac{2+\epsilon}{4-\epsilon})^m \ge (\frac{2+\epsilon}{4})^m = (\frac{1}{2} + \frac{\epsilon}{4})^m$$

(d) Show that $P(M) \ge P(G)P(M|G)$. Then show that

$$P(G) \le (1 + \frac{\epsilon}{2})^{-m}$$

The first inequality is pretty clear. It's easier for just M events to happen... and they'll happen more often than events happening in both M and G, even if there is an overlap.

But again, just definition chasing we get:

$$P(G) = \frac{P(M)P(G|M)}{P(M|G)}$$

Since we already have an inequality for $P(M|G)$, we apply it in reverse, since it's in the denominator. This looks like:

$$\frac{P(M)P(G|M)}{P(M|G)} \leq \frac{P(M)P(G|M)}{(\frac{1}{2} + \frac{\epsilon}{4})^m}$$

And we already know that $P(M) = p^m$, so we have:

$$P(G) \leq \frac{p^m * P(G|M)}{(\frac{1}{2} + \frac{\epsilon}{4})^m} = \frac{P(G|M)}{(1 + \frac{\epsilon}{2})^m}$$

And clearly $P(G|M)$ is upper bounded by 1, so we have:

$$P(G) \leq \frac{1}{(1 + \frac{\epsilon}{2})^m} = (1 + \frac{\epsilon}{2})^{-m}$$

(e) Combining the fact that for any $0 < \epsilon \leq 1$,

$$\ln(1 + \frac{\epsilon}{2}) > \frac{2}{5}\epsilon,$$

show that 2 (not pictured) holds.

Starting from the ln inequality:

$$\ln(1 + \frac{\epsilon}{2}) > \frac{2}{5}\epsilon$$

$$1 + \frac{\epsilon}{2} > e^{\frac{2}{5}\epsilon}$$

Since the larger denominator means a smaller number, this implies:

$$(1 + \frac{\epsilon}{2})^{-1} < (e^{\frac{2}{5}\epsilon})^{-1}$$

$$(1 + \frac{\epsilon}{2})^{-m} < (e^{\frac{2}{5}\epsilon})^{-m}$$

So then we have:

$$P(G) \leq (1 + \frac{\epsilon}{2})^{-m} < e^{-\frac{2}{5}\epsilon m}$$

Substituting $m = \frac{\epsilon n}{4}$:

$$P(G) \leq e^{-\frac{2}{5}\epsilon \frac{\epsilon n}{4}} = \exp(\frac{-\epsilon^2 n}{10})$$

(f) Compare (1) and (2) and argue why the Chernoff bound is better than the Chebyshev inequality.

For the same inequality the special case Chernoff bound gets exponentially closer to 0 as n grows, while the Chebyshev inequality goes linearly. In the "race to infinity" $\exp(\frac{-\epsilon^2 n}{10})$ grows **much** faster than $\epsilon^2 n$. And the tighter the bound, the more useful it is.