Probability Final

You may use your sheet from the midterm and a new sheet of notes. No calculators, cell phones, PDA's, laptops, or HAL 9000's. Show your reasoning. Don't just give the answer.

1. A box contains two gold balls and two clay balls. You are allowed to choose successively balls from the box at random. You win 1 dollar each time you draw a gold ball and lose 1 dollar each time you draw a clay ball. After a draw, the ball is not replaced.

   (a) If you draw exactly one ball, what is your expected earnings?{**answer:** $2/4 - 2/4 = 0$.}

   (b) What is the moment generating function for the value of the first draw?{**answer:** $(1/2)e^t + (1/2)e^{-t}$}

   (c) If you draw exactly k balls (for $k = 1, 2, 3, 4$) what is your expected earnings? {**answer:** 0}

   (d) If you draw until you are ahead by 1 dollar or until there are no more gold balls, what is your expected earnings?{**answer:** *GGcc = 1, GcGc = 1, GccG = 1, cGGc = 1, cGcG = 0, ccGG = 0. Grand total is 4, each equally likely, so 4/6, or 2/3.*}

2. Suppose you win 1 dollars when an black card is drawn from a deck of cards but you lose 1 dollar when a red card is drawn. (So out of the 52 cars, you win with 26 of them and lose with 26 of them.)

   (a) Let X_1 be the amount you win on the first draw, and X_2 be the amount you win on the second draw. (Assume you don't put the card back.) What is $Cov(X_1, X_2)$?

   {**answer:** $(26/52)(25/51) - (26/52)(26/51) - (26/52)(26/51) + (26/52)(25/51) = (1/2)(1/52)(25 - 26 - 26 + 25) = (1/2)(1/51)(-2) = 1/51$ }

   (b) What is the mean and variance of $X_1 + X_2$?{**answer:** *.25 - 2/51*}

   (c) What is the mean and variance of $\sum_{i=1}^{52} X_i$? (Hint: think before you compute.){**answer:** *0*}

3. Consider a non-negative random varible: $X \geq 0$.

   (a) If $E(X) = 1$, what is a good bound on $P(X \geq 100)$?

   (b) If $E(X) = 1$, and $V(X) = 1$ what is a good bound for $P(X \geq 100)$?

   (c) If the generating function $h_X(2) = 4$, (I.e. $E(2^X) = 4$ then what is a good bound for $P(X \geq 100)$?

4. Suppose the moment generating function for X is $g(t) = 1 + t$. In other words, $E(e^{tX}) = e^t$. What can you tell me about X? {**answer:** *$E(X) = 1$, $E(X^2) = 1$, so $var(X) = 0$ so $P(X = 1) = 1$.*}

5. The law of large numbers tells us alot about a sum of random variables. The CLT tells us even more about sums. But what about products? Let X_i be a random variable that takes on either $+1$ or -1 with equal probability. Let $P_n = \prod_{i=1}^{n} X_i$. Will P_n converge to some fixed value? (I.e. law of large numbers?) If it converges, what is this value, if it doesn't converge, what does P_n look like?{**answer:** *Does not converge. $P(P_n = 1) = P(P_n = -1) = .5$.*}

6. Statistics is often driven by two things, a prediction and a residual. Define the random variable $Z = E(Y|X)$ and the random variable $W = Y - Z$. Then Z is the prediction and W is the residual of the "regression" of Y on X.

   (a) What is $E(Z)$?{**answer:** $E(Z) = E(E(Y|X)) = E(Y)$}

   (b) What is $E(XZ)${**answer:** $E(XZ) = E(XE(Y|X)) = E(E(XY|X)) = E(XY)$}

   (c) Let $h()$ be an arbitary function, show $E(h(X)Z) = E(h(X)Y)$. {**answer:** $E(g(X)Z) = E(g(X)E(Y|X)) = E(E(g(X)Y|X)) = E(g(X)Y)$}

   (d) What is $E(WZ)$?{**answer:** $E(WZ) = E(E(WZ|X)) = E(ZE(W|X)) = E(ZE(Y - Z|X)) = E(ZE(Y|X) - ZE(Z|X)) = E(Z^2 - Z^2) = 0$}

7. Let X_i be a random variable with mean 1.01 and standard deviation .2. (For example, $X = 1.21$ or $X = .81$ with equal probability, but that is such an ugly statement, lets pretend I didn't mention it.) Let $W = \prod_{i=1}^{n} X_i$. Suppose all the X_i's are independent, so the whole series is IID.

   (a) What is $E(W)$?{**answer:** $E(W) = 1.01^n$}

   (b) What is the long run growth rate (i.e. $\lim_{n\to\infty}(\log W_n)/n$)?{**answer:** $\log W_n)/n \approx E(\log(X)) \approx E(X-1) - V(X)/2 = .01 - .2^2/2 = -.01$}

   (c) What will W_{1000} look like?{**answer:** $.99^{1000} = e^{-.01*1000} = e^{-10} \approx .0001$}

8. Suppose you put 100 mice on a calorie restriction diet. Normal mice on a normal diet live 1000 days with a standard deviation of 150 days.

   (a) If this diet doesn't change the length of life for these mice, what will be the mean, variance and distribution of $\overline{T}$? (Where $\overline{T}$ is the average number of days a mouse in the experiment lives.){**answer:** $E(\overline{T}) = 1000$, $V(\overline{T}) = (15)^2$, *it is normal.*}

   (b) Find a good estimate the probability that $\overline{T}$ is bigger than 1300.{**answer:** $(\overline{T} - 1000)/15 = 20$. *So via Chebeshev, the probability is less than* $1/20^2$, *or 1/400.*}

   (c) If your experiment actually yielded an average of 1300, would you believe that these mice have the same mean as typical mice?{**answer:** *Nope! Low feed mice live longer.*}

   (d) (bonus) Using generating functions, provide a better bound.{**answer:** *If you have a really big table for normals, it might go out to twenty, and so you know the answer is about 1/googol. (I mentioned this fact in class–but this is just an aside.) Here are your steps: (1) compute the moment generating function to be* $g(t) = e^{t^2/2}$. *(2) Now use Markov for* $P(e^{tZ} > e^{t20}) < e^{t^2/2}/e^{t20}$. *(3) Now optimize this over* t *to see that* $t = 20$ *gets the best bound. (4) the bound is now* e^{-200}.}