Problem Set 12

1. Total Annihilation
A super power has 2620 missiles stored in well separated silos. An enemy is considering a sneak attack. However, for the attack to succeed every one of the missiles must be destroyed. Assume that each attacking warhead hits one of the enemy missiles with each enemy missile being equally likely to be the one that is hit. How many warheads on the average will be needed to ensure the complete destruction of every enemy missile?

2. Significance Levels
A study has shown that in a certain profession the women are earning only 88% as much on the average, as their male counterparts earn. However, the study is several years old and a women's organization wants to determine if the women in this profession are losing relative to the men. It is known that the men in the profession now receive 138% of the pay they received when the above study was conducted. A random sample of female professionals is made. The average pay of these women was found to be the following (as a percentage of the current average pay of men in the profession): 70%, 78%, 80%, 83%, 84%, 86%, 87%, 96%.

   (a) Compute the sample mean and standard deviation. Note that if the sample average is m, then the sample variance is given by

   $$\frac{(x_1 - m)^2 + \cdots (x_n - m)^2}{n-1}$$

   To see why we divide by $n - 1$ rather than n see the extra credit problem.

   (b) At the .05 significance level are the women in this profession losing ground relative to the men? Does this coincide with your "gut feeling" in this problem?

3. Confidence Intervals
If a set of grades on the a discrete math examination are approximately normally distributed with a mean of 82 and a standard deviation of 6.9, find

   (a) The lowest passing grade if the lowest 5% of the students are given F's.

   (b) The highest B if the top 10% of the students are given A's.
   Hint: You may assume that if X is normal with mean 0 and variance 1, then $P[X \leq 1.3] = .9$ and $P[X \leq 1.65] = .95$.

4. Polya Urns
Recall from lecture that in a Polya Urn process with k urns and n balls, the minimum occupancy in an urn is expected to be about n/k^2 and the maximum about $n \log n/k$. In this question you will prove the first bound, by showing that the probability that the minimum occupancy is $\leq 2n/k^2$ is at least $1 - 1/e$.

   For convenience we will show this bound of $2n/k^2$ when the number of urns is $k + 1$ rather than k. Recall that the urn process corresponds to picking at random k balls (barriers) out of $n + k$ balls arranged in random order on a line. If any of these chosen barriers (balls) are within $2n/k^2$ of each other then we have an urn with low occupancy as desired.

   (a) Let us assume that the first $k/2$ chosen barriers (balls) satisfy that no two are within $2n/k^2$ of each other. What is the probability that if a new ball is picked it is not within $2n/k^2$ of one of these $k/2$ balls?

(b) Show that the probabibilty that none of the $k/2$ remaining balls lies within $2n/k^2$ of these $k/2$ balls is at most $1/e$. Show that this implies the result we are looking for.

5. Extra Credit
Suppose we have n samples $x_1 \dots x_n$ drawn from some distribution. We wish to estimate the mean and variance of this distribution. The obvious choice for an estimate of the mean m is $\frac{x_1+\cdots+x_n}{n}$. Prove that $E[m] = E[x_i]$.

Now the estimate for the variance is

$$v = \frac{(x_1 - m)^2 + \cdots (x_n - m)^2}{n - 1}$$

Prove that $E[v] = E[x_i^2] - E[x_i]^2$. i.e. the expected value of v is the variance of the distribution we are sampling from.