Statistics
Homework Two

Due approximately 1 February, 2008. I don't think many of these problems will entail long calculations. Feel free to make use of results from Wasserman or elsewhere, but be sure to state what results you're using. E.g., there's no need to derive the result for the mean of a Poisson distribution.

1. Problem one from Chapter 3 of Wasserman.

2. Let X be the number of cars that pass a certain point on Eagle Lake Road in fifteen minutes. Let X be distributed with a Poisson distribution with a mean of 6. This is a reasonably realistic assumption. At least the Poisson part. (The number 6 is fictional.) Knowing the distribution of cars along a road and also something about the frequency with which animals cross the road is important for modeling the distribution of roadkill and trying to come up with strategies to prevent roadkill.[1]

   (a) Sketch $f(x)$. Remember that X is discrete.

   (b) Sketch $F(x)$. Remember that X is discrete. So $F(x)$ will not be continuous.

   (c) What is the probability that there are 6 cars in a particular fifteen-minute time interval?

   (d) What is the probability that there are 15 cars in a particular fifteen-minute time interval?

   (e) What is $\mathbb{V}(X)$?

3. The number of errors, per page of text, is often Poisson distributed. Let G be the number of grammar errors per page and S the number of spelling errors per page. Let $S \sim \text{Poisson}(3)$ and $G \sim \text{Poisson}(1)$. Suppose you are penalized 4 points for every grammar error and 2 points for every spelling error. Let P denote the penalty per page.

   (a) Find $\mathbb{E}(G)$.

   (b) Find $\mathbb{E}(S)$.

   (c) Find $\mathbb{V}(G)$.

   (d) Find $\mathbb{V}(S)$.

[1] See, e.g.: Waller JS, Servheen C and Patterson DA. 2006. Probabilistic measure of road lethality. IN: Proceedings of the 2005 International Conference on Ecology and Transportation, Eds. Irwin CL, Garrett P, McDermott KP. Center for Transportation and the Environment, North Carolina State University, Raleigh, NC: pp. 503-508. `http://repositories.cdlib.org/jmie/roadeco/Waller2005a/`. See also, Szerlag S and McRobert SP. 2006. Road ecology of the northern diamondback terrapin, *Malaclemys terrapin terrapin*. IN: Proceedings of the 2005 International Conference on Ecology and Transportation, Eds. Irwin CL, Garrett P, McDermott KP. Center for Transportation and the Environment, North Carolina State University, Raleigh, NC: pp. 634-637. `http://repositories.cdlib.org/jmie/roadeco/Szerlag2005a/`. The latter paper concerns turtles in Southern New Jersey.

(e) Find $\mathbb{E}(P)$.

(f) Find $\mathbb{V}(P)$.

4. **I will soon be adding one additional problem on the exponential distribution.**

5. For these problems use the results on page 65 for confidence intervals for Bernoulli variables derived from Hoeffding's Inequality.

   (a) You are conducting a poll to determine the fraction of registered democrats who will vote for Hillary Clinton. Call this fraction p. You wish to estimate p to with 0.02 at the 95% confidence level. How many people do you need to include in your poll?

   (b) YOu flip a coin 500 times and observe 471 heads. What is the 95% confidence interval for your estimate of p, where p is the probability that the coin comes up heads? What is the 99% confidence interval?

6. **Optional!** Consider a game of chance that is defined as follows. You flip a fair coin until the outcome is H. Let L equal the number of tosses needed until you see an H. For example, if you observed TTH, then $L = 3$. Your payoff M is given by $M = 2^L$. So if you toss TTH, you would get a payoff of 8. Show that $\mathbb{E}(M)$ does not exist.