Some BONUS review exercises due Wednesday 7-14-10. Students of section 201 had to miss class Friday last. Take this as a tutorial to prepare you also on some of the ideas from Friday.

1. For the list of x scores {2.3, 5.4, 7.9, 6.6, 2.1} determine

a. μ

b. σ^2

c. σ

d. $\sqrt{\left(\mu_{x^2} - \mu_x^2\right)}$ (i.e. first calculate the mean of squares then subtract from it the square of means) and compare with (**c**).

e. Sample standard deviation "s" calculated as either of

$$s = \sqrt{\frac{\sum(x-\mu)^2}{n-1}} = \sqrt{\frac{n}{n-1}}\ \sigma$$

2. The per-play mean and standard deviation of the return from a casino game can tell you a lot about what will likely happen when it is played many times. From the casino's point of view there are many plays. Suppose that no single play of a particular casino game has net casino returns of a draw from the list

x= {1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, -2, 2, 2, 2, 2, 2, 2, 2, -5, -10, -20}

We find

μ = 0.411765 and σ = 4.2573.

A player can be comfortable with this game because although they will most often lose 1 or 2 these losses will occasionally be punctuated by a win of 2, 5, 10 or 20. Here are five simulated examples of 10 plays each (they are individual plays, not accumulated, illustrating this fact:

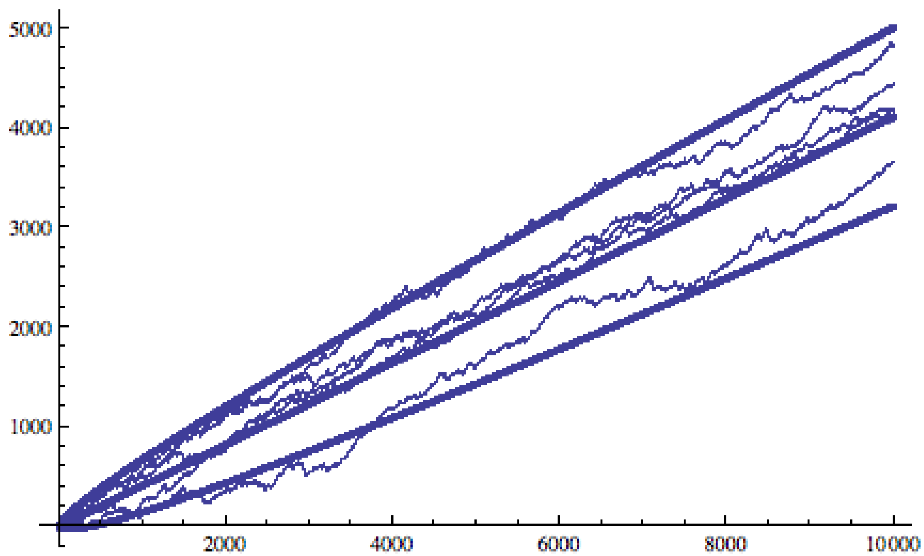
{2, 2, 1, -5, 1, 2, 1, 1, 1, -20}
{2, 2, 1, 2, 1, 2, 2, 2, 2, 2}
{1, 2, 2, -10, 1, 2, 2, -20, 2, 2}
{-20, 2, 2, 1, 2, 2, 2, 2, -2, 1}
{2, 1, 2, 1, 2, 2, 2, 2, 2, 2}

This is enough to keep players interested, especially since any player wins (like the -20, which means the player won 20) are witnessed by many other players as well, reinforcing the idea that it is possible to win.

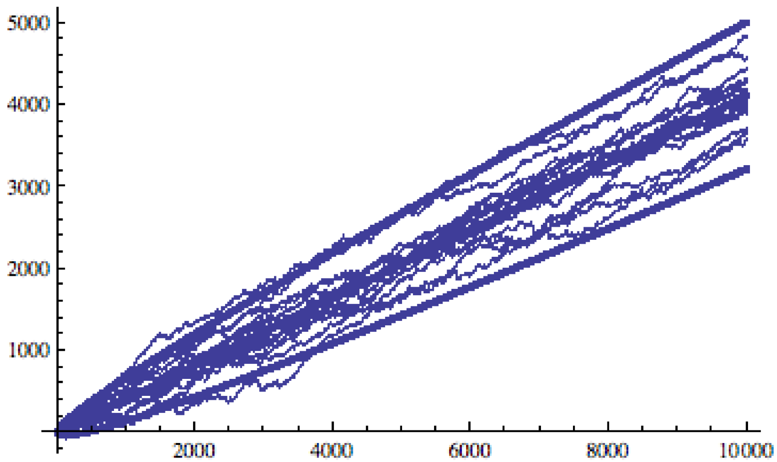
THE ENVELOPE:

3. Refer to problem 2. Over the long run the casino does well. Here are five unrelated plots of the casino's play by play **accumulated** returns from 10000 plays. Each is shown in a plot of the likely "long run" envelope for the behavior of a running total accumulated by the casino over the course of 10000 plays. Early behavior may fall outside the envelope but later behavior should settle into the envelope or fall negligibly outside of it.

Plays are made **independently** (it means no plays influence the results of other plays).



Here are fifteen such runs of 10000. See how it begins to fill up the ideal parabolic region around the trend line?



So all but a negligible fraction of 10000-length plots of the random accumulations of casino returns

* Respect the trend line of μ per play average return.
* Respect (after a while) the theoretical envelope:

$$\text{upper} = \mu\,\# + \sigma \sqrt{2\,(\sharp\ \text{Log}(\text{Log}(\sharp))}$$

$$\text{lower} = \mu\,\# - \sigma \sqrt{2\,(\sharp\ \text{Log}(\text{Log}(\sharp))}$$

as the number of plays (denoted by "#") varies from 1 to 10000. Moreover, collectively a great many such 10000 length plots more or less fill up and define the region (except at the beginning or left side of the plot it may be ragged and not so close to the theoretical ideal represented b the upper and lover parts of the envelope).

a. Use your calculator to check that the upper and lower boundaries defined mathematically agree at several values of # (in 1 to 10000) with those plotted. Log(Log #)) is the result of taking a number such as # = 5,000, taking its $\log_e$ (the natural log key "ln" perhaps) and then the log of that:

$$\log(\log 5000)) \sim \log(8.51719) \sim 2.14209$$

So the upper boundary at # = 5,000 is around

$$\mu\,\# + \sigma \sqrt{2\,(\sharp\ \text{Log}(\text{Log}(\sharp))} \sim$$

$$\sim 0.411765\ \ 5000 + 4.2573 \sqrt{2\ \ 5000\ \ \text{Log}(\text{Log}(5000)}$$

~ 2682 (so check that it is so in the plot).

b. Notice that for # = 10000 by itself the plots seem to finish closer to the trend value 0.411765 5000 ~ 2059 than the envelope allows for. That is because the envelopes have to be wide enough to include most of all paths all along the way. So they have to be wider than to include only most end results for # = 10000.

SNAPSHOT AT A PARTICULAR LARGE # OF PLAYS.

4. Refer to problem 2. For any fixed large # the total casino winnings up to that point in time approximately follow a normal (i.e. "Bell") curve with

bell curve mean = μ #

bell curve standard deviation = σ $\sqrt{\#}$

From Key 6 "empirical rule"

Around 68% of plots of 10000 plays end up in the range

$0.411765 \; 10000 \pm 4.2573 \; \sqrt{10000}$.

Around 95% of plots of 10000 plays end up in the range

$0.411765 \; 10000 \pm 1.96 \; 4.2573 \; \sqrt{10000}$.

a. Evaluate the 68% interval given above.

b. What is the chance the casino earns a total of less than

$0.411765 \; 10000 \sim 4117$

after 10000 plays?

c. What is the chance the casino earns a total of less than

$0.411765 \; 10000 - 1.96 \; 4.2573 \; \sqrt{10000} \sim 3283$

after 10000 plays? Hint: This is the left end of a 95% interval.

d. Compare the size of the amount the casino is looking to earn

$0.411765 \; 10000 \sim 4117.65$

with the interval $0.411765 \; 10000 \pm 1.96 \; 4.2573 \; \sqrt{10000}$ within which the actual random return falls around 95% of the time. By what percentage might actual return differ from projected return and still be within this 95% range?