Stat 100a: Introduction to Probability.

Outline for the day:

1. HW and exams.
2. More on luck and skill.
3. Notes on pdfs, survivor functions, E(X+Y), and independence.
4. Cards til 2^{nd} king.
5. Pareto random variables.
6. Covariance and correlation.
7. Notes on hw3.
8. Testing out your R function.
9. Conditional expectation.
10. Law of Large Numbers.
11. Central Limit Theorem.

Read through chapter 7.3.

1. HW and exams.

HAND IN HW2!

HW3 is on the course website now. It is due Thu Jul 23, 10am in class.
6.6, 6.10, 6.14, 7.2, 7.8, 7.14.

I will pass back exam 1 in alphabetical order by last name. Please be silent until I am finished passing them out.
The mean was 90% and the SD was 20%.

Exam 2 is Tue Jul 21, one week from today. All the exams are cumulative.

2. More on luck and skill.

Gain due to skill on a betting round = your expected *profit* during the betting round
= your exp. chips after betting round – your exp. chips before betting round
= (equity after round + leftover chips) –
(equity before round + leftover chips + chips you put in during round)
= equity after round – equity before round – cost during round.

For example, suppose you have A♣A♠, I have 3♥ 3♦, the board is
A♥ Q♣ 10♦ and there is $10 in the pot. The turn is 3♠.
You go all in for $5 and I call. How much equity due to skill did you gain on the turn?
Your prob. of winning is 43/44.
Your skill gain on turn = your equity after turn bets - equity before turn bets – cost
= ($20)(43/44) - ($10)(43/44) - $5
= $4.77.

Suppose instead you bet $5 on the turn and I folded. How much equity due to skill did you gain on the turn?
$15(100%) - ($10)(43/44) - $5 = $0.23.

3. Notes on pdfs, survivor functions, E(X+Y), and independence.

For some problems such as 6.10, it can be convenient to look at the survivor function or cdf first, and then get the pdf by noting that for any continuous RV, $F(x) = P(X \leq c) = \int_{-\infty}^{x} f(y)dy$, so $\mathbf{F'(x) = f(x)}$. This is the fund. thm. of calculus. And $F(x) = 1 - S(x)$. $S(x) \to F(x) \to f(x)$.

If X and Y are independent random variables, then
$E[f(X)\, g(Y)] = E[f(X)]\, E[g(Y)]$, for any functions f and g.
See Exercise 7.12. This is useful for exercise 5.4 for instance.

Also, recall the fact that $E(X+Y) = E(X) + E(Y)$, even if X and Y are dependent.

4. Cards til 2^{nd} king.

Deal the cards face up, without reshuffling.

Let Z = the number of cards til the 2nd king.

What is E(Z)?

The solution uses the fact

$$E(X+Y+Z + ...) = E(X) + E(Y) + E(Z) + ...$$

♣ ♥ ◆ ♠ ♣ ♥ ◆ ♠

E(cards til 2^{nd} king).

Z = the number of cards til the 2nd king. What is E(Z)?

Let X_1 = number of non-king cards before 1^{st} king.

Let X_2 = number of non-kings after 1^{st} king til 2^{nd} king.

Let X_3 = number of non-kings after 2^{nd} king til 3^{rd} king.

Let X_4 = number of non-kings after 3^{rd} king til 4^{th} king.

Let X_5 = number of non-kings after 4^{th} king til the end of the deck.

Clearly, $X_1 + X_2 + X_3 + X_4 + X_5 + 4$ always = 52.

So $X_1+X_2+X_3+X_4+X_5 = 48$. Therefore $E(X_1)+E(X_2)+E(X_3)+E(X_4)+E(X_5) = 48$.

By symmetry, $E(X_1) = E(X_2) = E(X_3) = E(X_4) = E(X_5)$.

Therefore, $E(X_1) = E(X_2) = 48/5$.

$Z = X_1+X_2+2$, so $E(Z) = E(X_1) + E(X_2) + 2 = 48/5 + 48/5 + 2 = 21.2$.

5. Pareto random variables. ch6.6

Pareto random variables are an example of *heavy-tailed* random variables, which means they have very, very large outliers much more frequently than other distributions like the normal or exponential.

For a Pareto random variable, the pdf is $f(y) = (b/a)\,(a/y)^{b+1}$, and the cdf is

$F(y) = 1 - (a/y)^b$,

for $y>a$, where $a>0$ is the *lower truncation point*, and $b>0$ is a parameter called the *fractal dimension*.

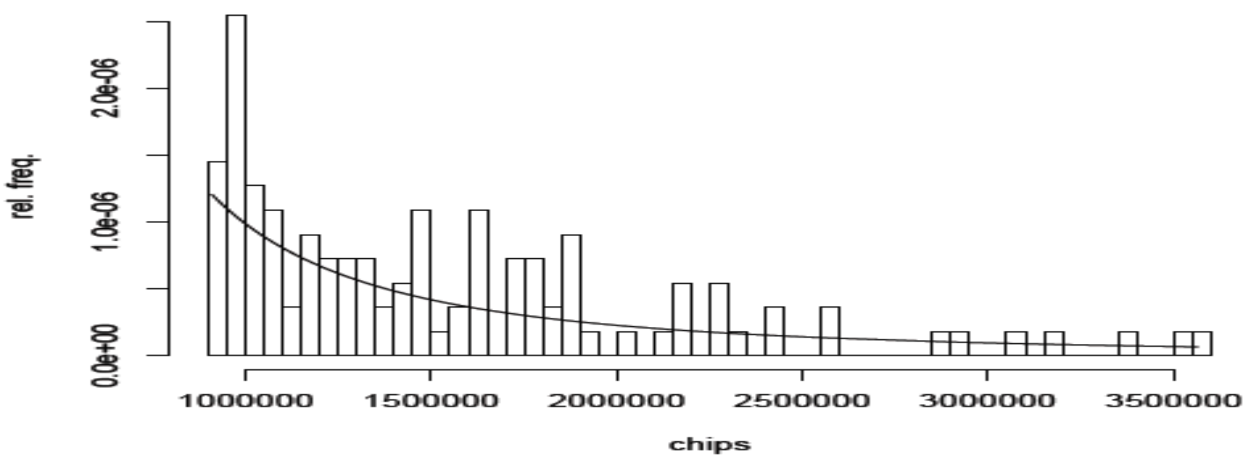


Figure 6.6.1: Relative frequency histogram of the chip counts of the leading 110 players in the 2010 WSOP main event after day 5. The curve is the Pareto density with $a = 900{,}000$ and $b = 1.11$.

6. Covariance and correlation, p127.

For any random variables X and Y,

$\mathrm{var}(X+Y) = E[(X+Y)]^2 - [E(X) + E(Y)]^2$

$= E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 + 2E(XY) - 2E(X)E(Y)$

$= \mathrm{var}(X) + \mathrm{var}(Y) + 2[E(XY) - E(X)E(Y)]$.

$\mathrm{cov}(X,Y) = E(XY) - E(X)E(Y)$ is called the *covariance* between X and Y,

$\mathrm{cor}(X,Y) = \mathrm{cov}(X,Y) / [SD(X)\ SD(Y)]$ is called the *correlation* bet. X and Y.

If X and Y are ind., then $E(XY) = E(X)E(Y)$,

so $\mathrm{cov}(X,Y) = 0$, and $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$.

Just as $E(aX + b) = aE(X) + b$, for any real numbers a and b,

$\mathrm{cov}(aX + b,Y) = E[(aX+b)Y] - E(aX+b)E(Y)$

$= aE(XY) + bE(Y) - [aE(X)E(Y) + bE(Y)] = a\ \mathrm{cov}(X,Y)$.

Ex. 7.1.3 is worth reading.

X = the # of 1^{st} card, and Y = X if 2^{nd} is red, -X if black.

$E(X)E(Y) = (8)(0)$.

$P(X = 2 \text{ and } Y = 2) = 1/13 * ½ = 1/26$, for instance, and same with any other combination,

so $E(XY) = 1/26\ [(2)(2)+(2)(-2)+(3)(3)+(3)(-3) + ... + (14)(14) + (14)(-14)]$

$= 0$.

So X and Y are *uncorrelated*, i.e. $\mathrm{cor}(X,Y) = 0$.

But X and Y are not independent.

$P(X=2 \text{ and } Y=14) = 0$, but $P(X=2)P(Y=14) = (1/13)(1/26)$.

7. HW3 notes.

Problem 7.14 refers to Theorem 7.6.8, p152.

You have k of the n chips in play. Each hand, you gain 1 with prob. p, or lose 1 with prob. $q=1-p$.

Suppose $0<p<1$ and $p\neq 0.5$. Let $r = q/p$.

Then P(you win the tournament) = $\mathbf{(1-r^k)/(1-r^n)}$.

The proof is by induction, and is similar to the proof of Theorem 7.6.6.

Notice that if $k = 0$, then $(1-r^k)/(1-r^n) = 0$. If $k = n$, then $(1-r^k)/(1-r^n) = 1$.

For 7.14, the key things to remember are

1. If $p\neq 0.5$, then by Theorem 7.6.8, P(win tournament) = $(1-r^k)/(1-r^n)$, where $r = q/p$.
2. Let $x = r^2$. If $-x^3 + 2x - 1 = 0$, that means $(x-1)(-x^2 - x + 1) = 0$. There are 3 solutions to this.
   One is $x = 1$. The others occur when $x^2 + x - 1 = 0$, so $x = [-1 +/- \text{sqrt}(1+4)]/2 = -1.618$ or 0.618.
   So, $x = -1.618, 0.618$, or 1. Two of these possibilities can be ruled out.
   Remember that $p\neq 0.5$.

8. TESTING OUT YOUR FUNCTION FOR THE PROJECT.

Suppose your function is called “neverfold”. Run 6 neverfolds against 6 zeldas.

install.packages(holdem) ## you must be connected to the internet for this to work.

library(holdem)

a = neverfold

v = vera

decision1 = c(a,a,a,a,a,a,v,v,v,v,v,v)

name1 = c("n1","n2","n3","n4","n5","n6","v1","v2","v3","v4","v5","v6")

tourn1(name1, decision1) ## Do this line a few times. Make sure there’s no error.

9. Conditional expectation, E(Y | X), ch. 7.2.
Suppose X and Y are discrete.
Then E(Y | X=j) is defined as $\sum_k$ k P(Y = k | X = j), just as you'd think.
E(Y | X) is a **random variable** such that E(Y | X) = E(Y | X=j) whenever X = j.

For example, let X = the # of spades in your hand, and Y = the # of clubs in your hand.
a) What's E(Y)? **b)** What's E(Y|X)? **c)** What's P(E(Y|X) = 1/3)?

a. E(Y) = 0P(Y=0) + 1P(Y=1) + 2P(Y=2)
= 0 + 13x39/C(52,2) + 2 C(13,2)/C(52,2) = 0.5.

b. X is either 0, 1, or 2. If X = 0, then E(Y|X) = E(Y | X=0) and
E(Y | X=0) = 0 P(Y=0 | X=0) + 1 P(Y=1 | X=0) + 2 P(Y=2 | X = 0)
= 0 + 13x26/C(39,2) + 2 C(13,2) / C(39,2) = **2/3**.
E(Y | X=1) = 0 P(Y=0 | X=1) + 1 P(Y=1 | X=1) + 2 P(Y=2 | X = 1)
= 0 + 13/39 + 2 (0) = **1/3**.
E(Y | X=2) = 0 P(Y=0 | X=2) + 1 P(Y=1 | X=2) + 2 P(Y=2 | X = 2)
= 0 + 1 (0) + 2(0) = **0**.
So E(Y | X = 0) = 2/3, E(Y | X = 1) = 1/3, and E(Y | X = 2) = 0. That's what E(Y|X) is
c. P(E(Y|X) = 1/3) is just P(X=1) = 13x39/C(52,2) ~ 38.24%.

10. Law of Large Numbers (LLN) and the Fundamental Theorem of Poker, ch 7.3.

David Sklansky, *The Theory of Poker,* 1987.

> "Every time you play a hand differently from the way you would have played it if you could see all your opponents' cards, they gain; and every time you play your hand the same way you would have played it if you could see all their cards, they lose. Conversely, every time opponents play their hands differently from the way they would have if they could see all your cards, you gain; and every time they play their hands the same way they would have played if they could see all your cards, you lose."

Meaning?

LLN: If X_1, X_2 , etc. are iid with expected value μ and sd σ, then $\overline{X}_n$ ---> μ.

Any short term good or bad luck will ultimately become *negligible* to the sample mean.

However, this does not mean that good luck and bad luck will ultimately cancel out.

See p132.

11. The Central Limit Theorem (CLT), ch 7.4.

Sample mean $\overline{X_n} = \sum X_i / n$

iid: independent and identically distributed.

Suppose X_1, X_2 , etc. are iid with expected value μ and sd σ ,

LAW OF LARGE NUMBERS (LLN):

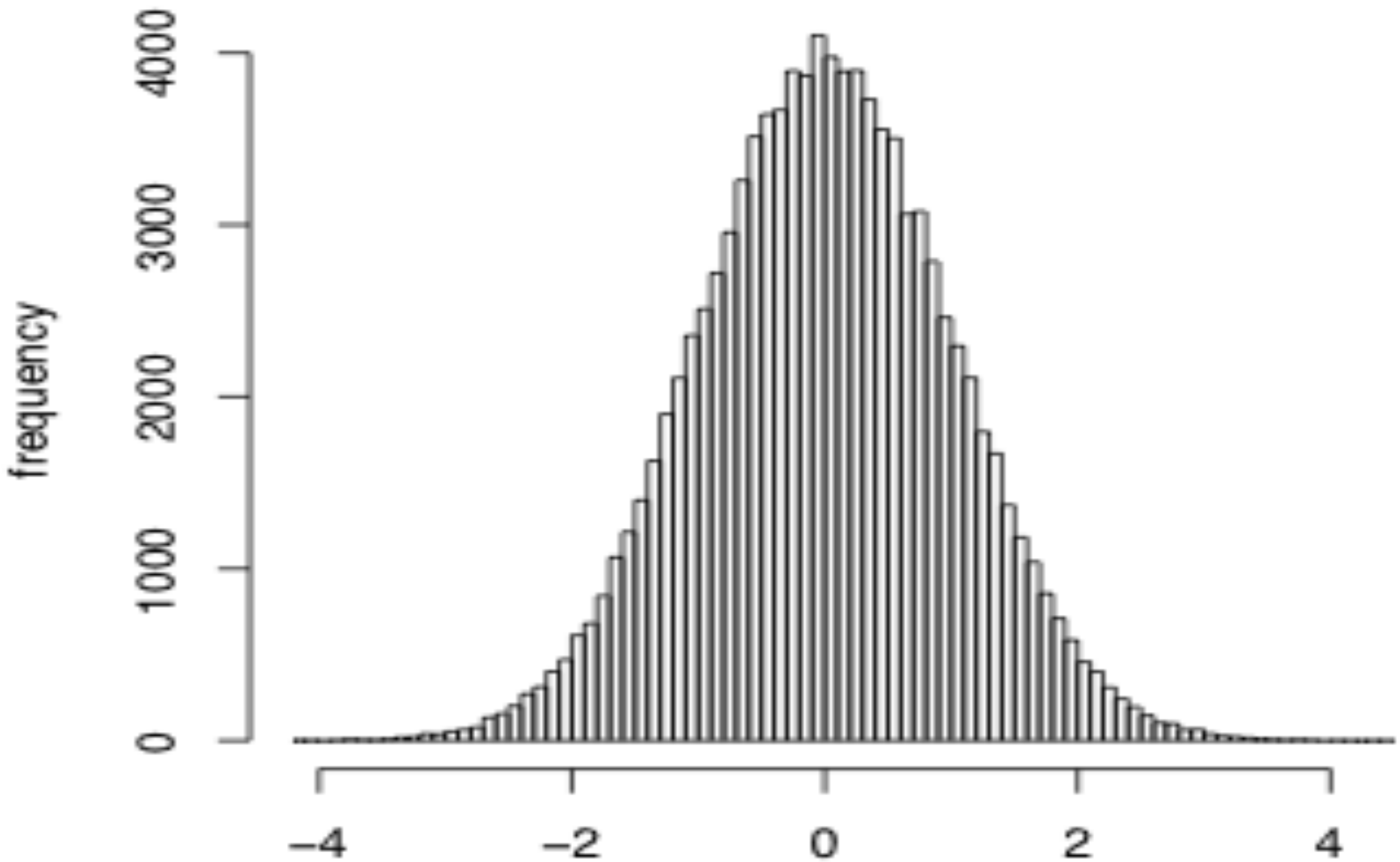
$\overline{X_n} \text{ ---> } \mu$.

CENTRAL LIMIT THEOREM (CLT):

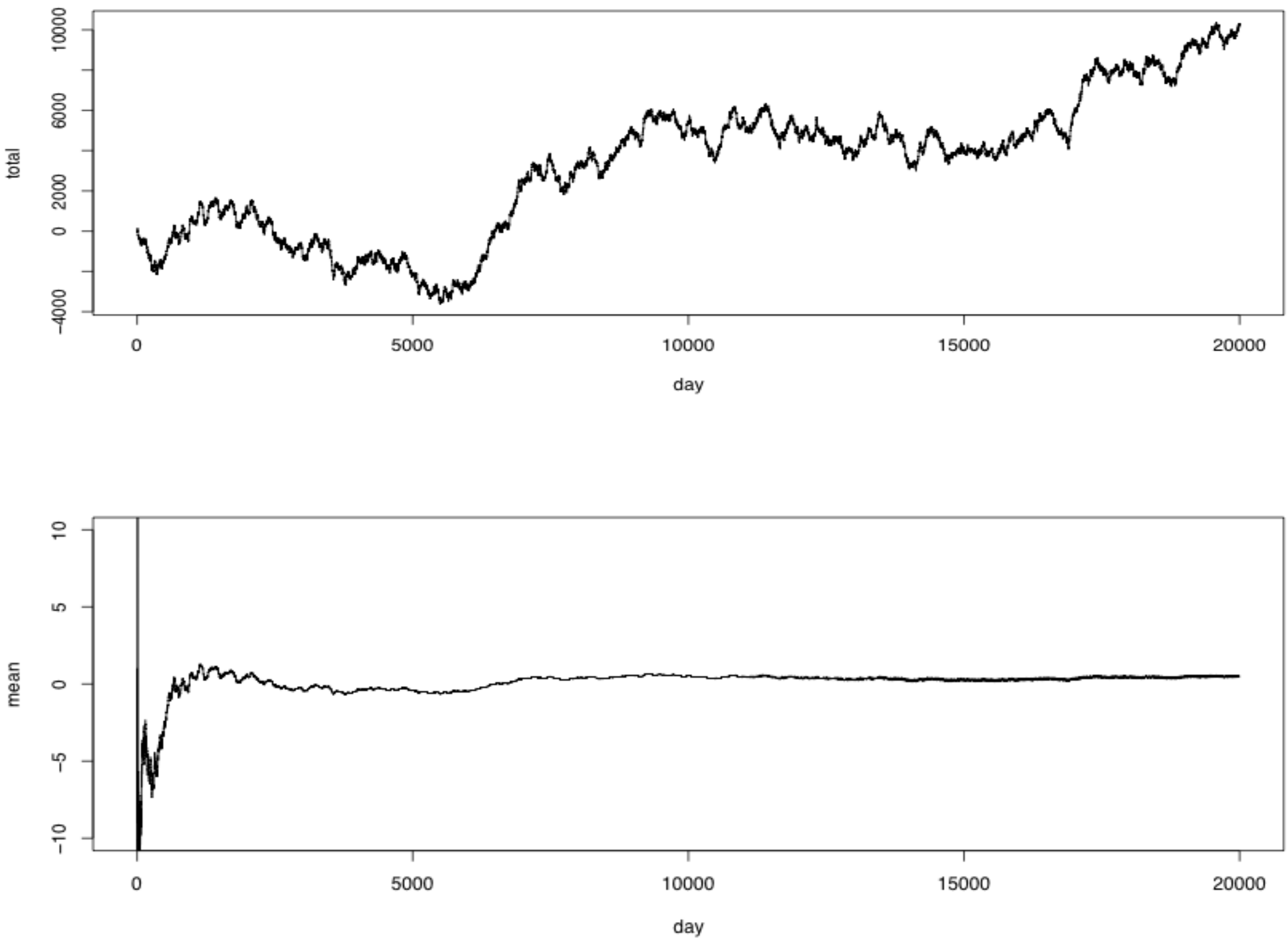
$(\overline{X_n} - \mu) \div (\sigma/\sqrt{n})$ ---> Standard Normal.

Useful for tracking results.

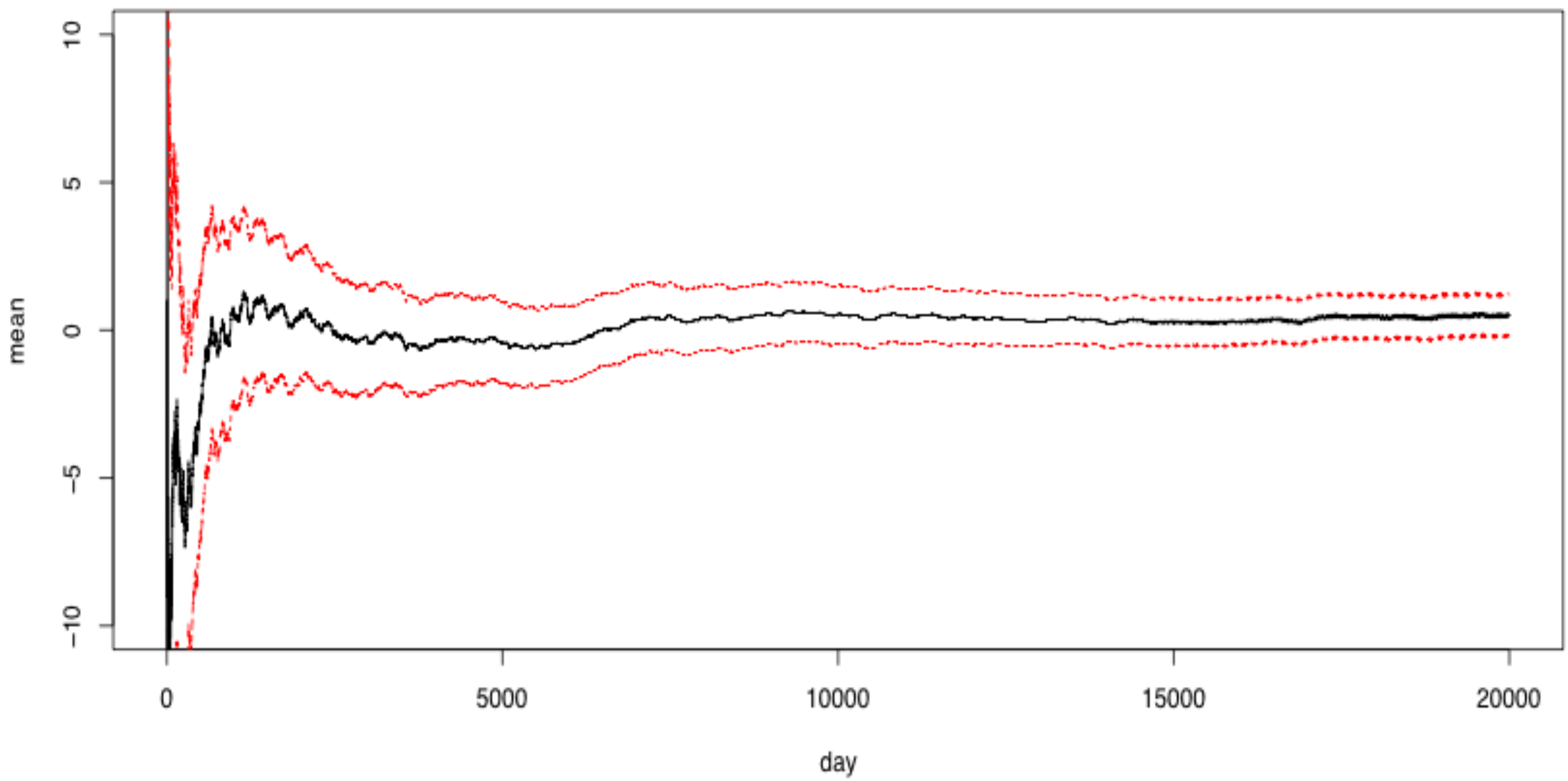
95% between -1.96 and 1.96



Truth: -49 to 51, exp. value $\mu = 1.0$



Truth: uniform on -49 to 51. $\mu = 1.0$
Estimated using $\overline{X}_n$ +/- 1.96 $\sigma/\sqrt{n}$
= .95 +/- 0.28 in this example



Central Limit Theorem (CLT): if X_1, X_2 …, X_n are iid with mean μ & SD σ, then

$$(\overline{X}_n - \mu) \div (\sigma/\sqrt{n}) \text{ ---> Standard Normal. (mean 0, SD 1).}$$

In other words, $\overline{X}_n$ has mean μ and a standard deviation of $\sigma \div \sqrt{n}$.

Two interesting things about this:

(i) As n --> ∞, $\overline{X}_n$ --> ***normal.*** Even if X_i are far from normal.

e.g. *average* number of pairs per hand, out of n hands. X_i are 0-1 (Bernoulli).

μ = p = P(pair) = 3/51 = 5.88%. $\sigma = \sqrt{(pq)} = \sqrt{(5.88\% \times 94.12\%)}$ = 23.525%.

(ii) We can use this to find **a range** where $\overline{X}_n$ is likely to be.

About 95% of the time, a std normal random variable is within -1.96 to +1.96.

So 95% of the time, $(\overline{X}_n - \mu) \div (\sigma/\sqrt{n})$ is within -1.96 to +1.96.

So 95% of the time, $(\overline{X}_n - \mu)$ is within $-1.96\,(\sigma/\sqrt{n})$ to $+1.96\,(\sigma/\sqrt{n})$.

So 95% of the time, $\overline{X}_n$ is within $\mu - 1.96\,(\sigma/\sqrt{n})$ to $\mu + 1.96\,(\sigma/\sqrt{n})$.

That is, 95% of the time, $\overline{X}_n$ is in the interval μ +/- 1.96 $(\sigma/\sqrt{n})$.

= 5.88% +/- 1.96(23.525%/$\sqrt{n}$). For n = 1000, this is 5.88% +/- 1.458%.

For n = 1,000,000 get 5.88% +/- 0.0461%.

Another CLT Example

Central Limit Theorem (CLT): if $X_1, X_2, \ldots, X_n$ are iid with mean μ & SD σ, then

$$(\overline{X}_n - \mu) \div (\sigma/\sqrt{n}) \text{ ---> Standard Normal. (mean 0, SD 1).}$$

In other words, $\overline{X}_n$ is like a draw from a normal distribution

with mean μ and standard deviation of $\sigma \div \sqrt{n}$.

That is, 95% of the time, $\overline{X}_n$ is in the interval μ +/- 1.96 $(\sigma/\sqrt{n})$.

Q. Suppose you average \$5 profit per hour, with a SD of \$60 per hour. If you play 1600 hours, let Y be your average profit over those 1600 hours. What is range where Y is 95% likely to fall?

A. We want μ +/- 1.96 $(\sigma/\sqrt{n})$, where μ = \$5, σ = \$60, and n=1600. So the answer is

\$5 +/- 1.96 x \$60 / $\sqrt{(1600)}$

= \$5 +/- \$2.94, or the range [\$2.06, \$7.94].