Stat 100a: Introduction to Probability.

Outline for the day:

1. CLT.
2. CIs.
3. Sample size calculations.
4. Gold vs. Farha.
5. P(flop an ace high flush), P(flop a straight flush), etc.
6. Time til suited king.
7. Hellmuth vs. Farha, P(flush | suited cards).

Midterm is on chapters 1-7.5, except chapter 6.7 which we are skipping and the bulk of 6.3 about optimal play in the case of uniform hands, which we are also skipping.

♣ ♥ ◆ ♠

1. 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):

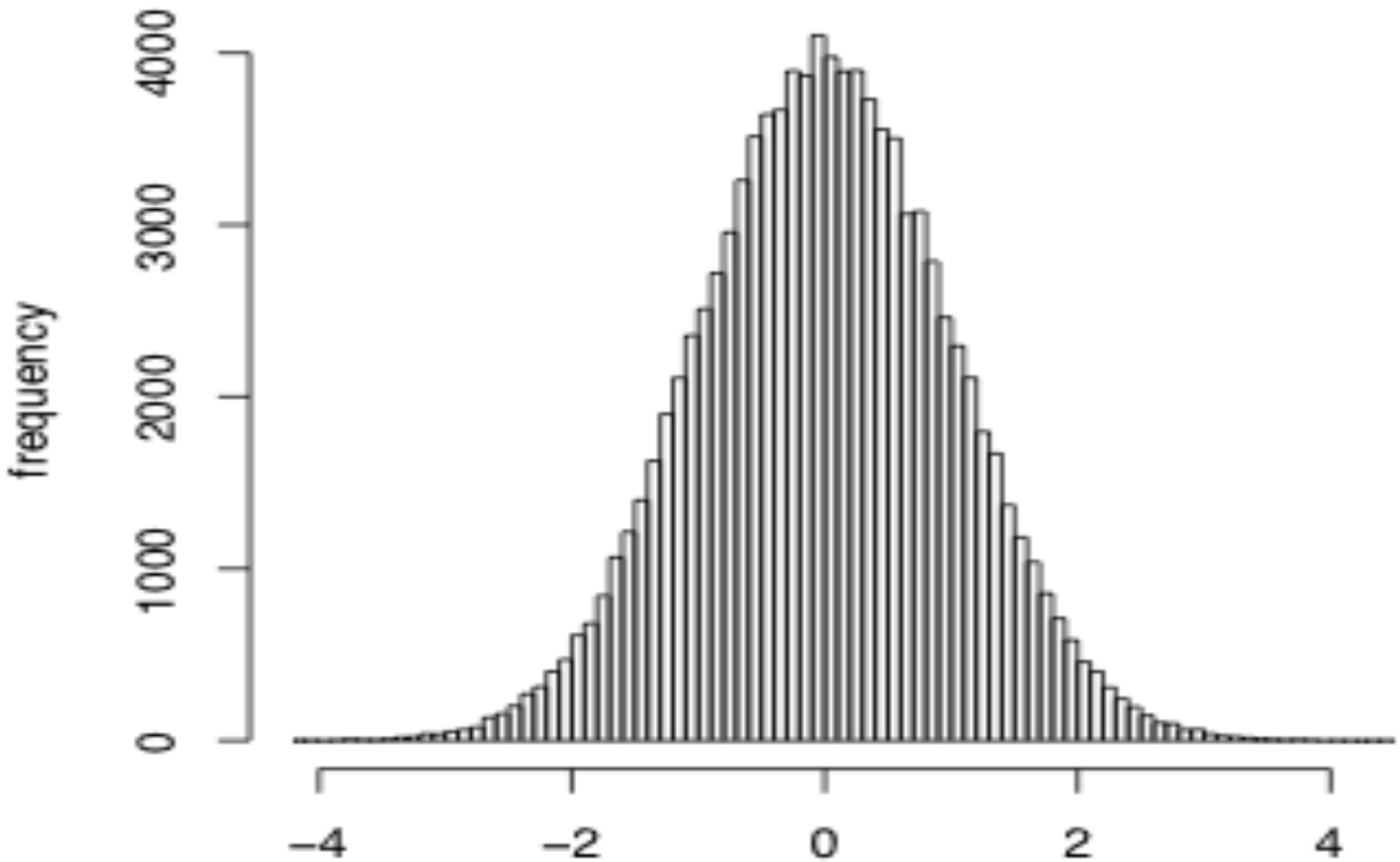
$\overline{X_n}$ ---> μ .

CENTRAL LIMIT THEOREM (CLT):

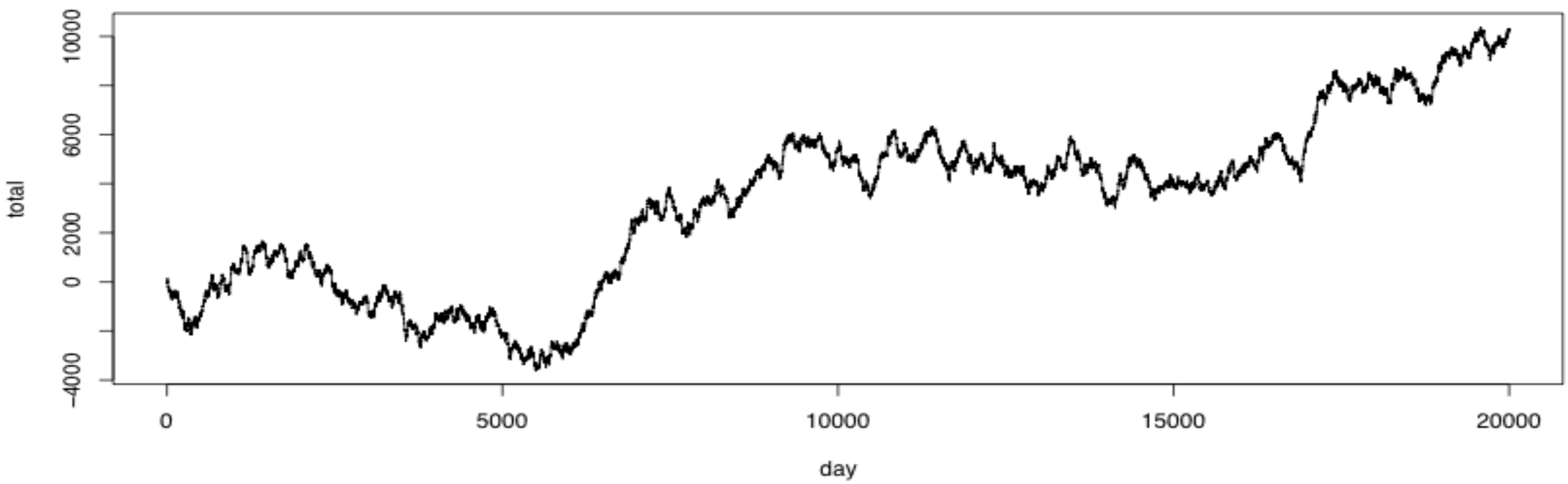
$(\overline{X_n} - \mu) \div (\sigma/\sqrt{n})$ ---> Standard Normal.

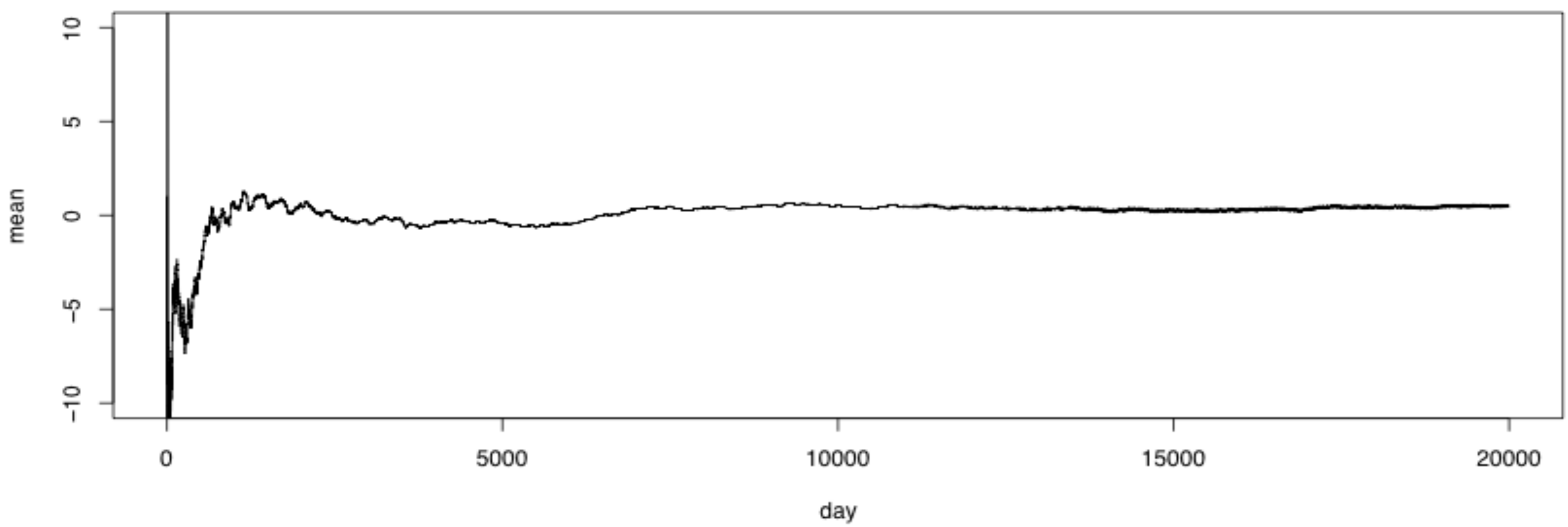
Useful for tracking results.

95% between -1.96 and 1.96

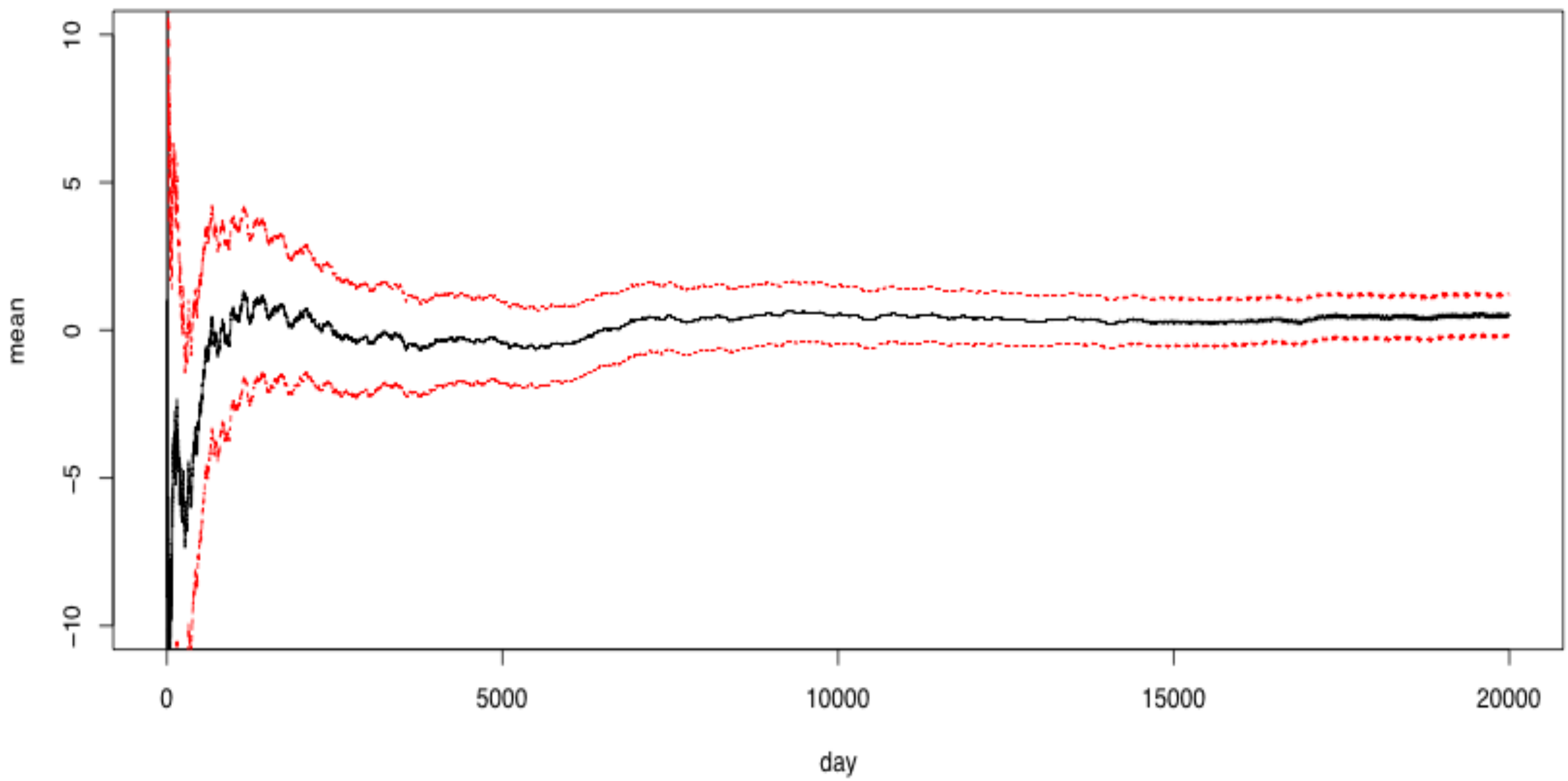


Truth: -49 to 51, exp. value $\mu = 1.0$





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



<u>Central Limit Theorem (CLT):</u> 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, $\overline{X}_n$ 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].

2. Confidence Intervals (CIs) for μ, ch 7.5.

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

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

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

Typically you know $\overline{X}_n$ but not μ. Turning the blue statement above around a bit means that 95% of the time, μ is in the interval $\overline{X}_n$ +/- $1.96\ (\sigma/\sqrt{n})$.

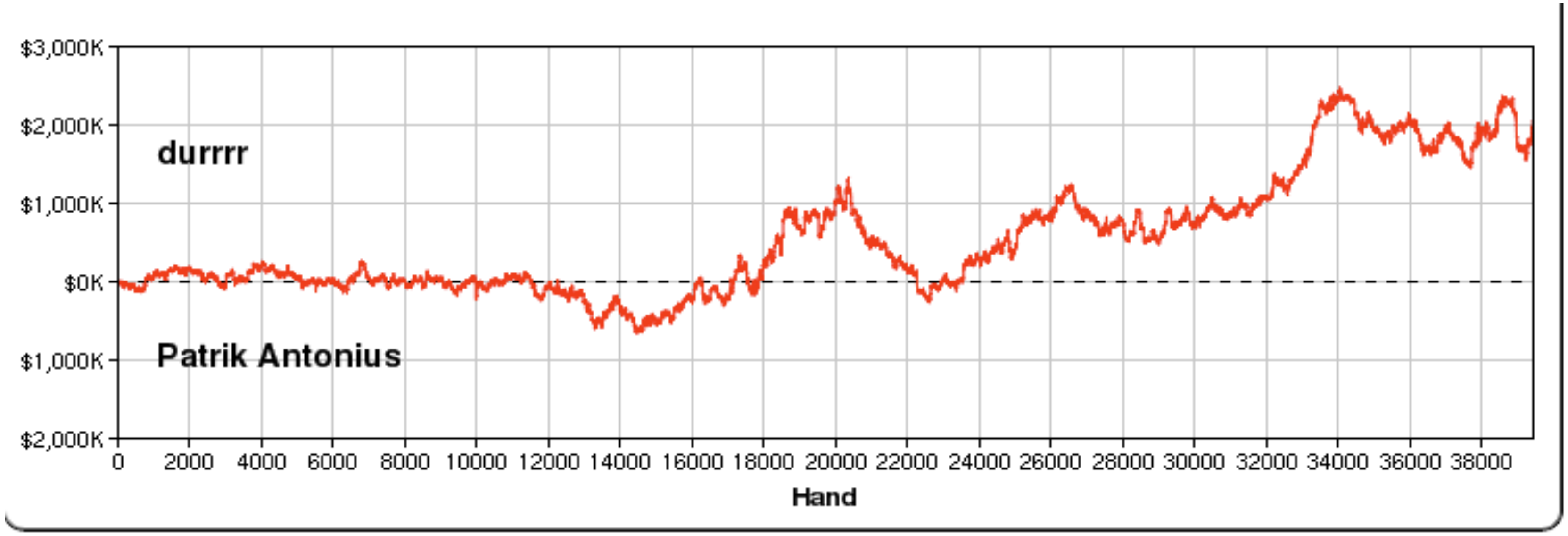
This range $\overline{X}_n$ +/- $1.96\ (\sigma/\sqrt{n})$ is called a 95% confidence interval (CI) for μ.

[Usually you don't know σ and have to estimate it using the sample std deviation, s, of your data, and $(\overline{X}_n - \mu) \div (s/\sqrt{n})$ has a t_{n-1} distribution if the X_i are normal. For n>30, t_{n-1} is so similar to normal though.]

$1.96\ (\sigma/\sqrt{n})$ is called the *margin of error.*

The range $\overline{X}_n$ +/- 1.96 (σ/√n) is a 95% confidence interval for μ. 1.96 (σ/√n)

(from fulltiltpoker.com:)



Based on the data, can we conclude Dwan is a better player? Is his longterm avg. $\mu > 0$?

Over these 39,000 hands, Dwan profited $2 million. $51/hand. sd ~ $10,000.

95% CI for μ is $51 +/- 1.96 ($10,000 / √39,000) = $51 +/- $99 = (-$48, $150).

Results are inconclusive, even after 39,000 hands!

3. Sample size calculation. How many *more* hands are needed?

If Dwan keeps winning $51/hand, then we want n so that the margin of error = $51.

1.96 (σ/√n) = $51 means 1.96 ($10,000) / √n = $51, so n = [(1.96)($10,000)/($51)]^2 ~ 148,000, so about 109,000 *more* hands.

4. Gold vs. Farha. Gold: 10♦ 7♠ Farha: Q♦ Q♠ Flop: 9♦ 8♥ 7♣

a) Who really is the favorite (ignoring all other players' cards)?

Gold's outs: J, 6, 10, 7. (4 + 4 + 3 + 2 = 13 outs, 32 non-outs)

P(Gold wins) = P(Out Out but not JT or Jx or 6x or 10y [y≠Q,9,8] or 7z [z≠Q])

= [choose(13,2) - 4*3 + 4*32 + 4*32 + 3*24 + 2*30] ÷ choose(45,2)

= 454 ÷ 990 = 45.86%.

b) What would you guess Gold had?

Say he'd do that 50% of the time with a draw,

100% of the time with an overpair,

and 90% of the time with two pairs. (and that's it)

Using Bayes' rule, P(Gold has a DRAW | Gold raises ALL-IN)

$$= \frac{[P(\text{all-in} \mid \text{draw}) * P(\text{draw})]}{[P(\text{all-in} \mid \text{draw})\, P(\text{draw})] + [P(\text{all-in} \mid \text{overpair})\, P(\text{overpair})] + [P(\text{all-in} \mid \text{2pairs})\, P(\text{2 pairs})]}$$

= [50% * P(draw)] ÷ [50% * P(draw)] + [100% * P(overpair)] + [90% * P(2 pairs)]

5. P(flop an ace high flush)? [where the ace might be on the board]

-- 4 suits

-- one of the cards must be an ace. choose(12,4) possibilities for the others.

So P(flop ace high flush) = 4 * choose(12,4) / choose(52,5)

= 0.0762%, or 1 in **1313**.

P(flop a straight flush)?

-- 4 suits

-- 10 different straight-flushes in each suit. (5 high, 6 high, …, Ace high)

So P(flop straight flush) = 4 * 10 / choose(52,5)

= 0.00154%, or 1 in **64974**.

6. What is the probability that you will be dealt a king and another card of the same suit as the king?

$4 * 12 / C(52,2) = 3.62\%$.

The typical number of hands until this occurs is ...

$1/.0362 \sim 27.6$.
$(\sqrt{96.38\%}) / 3.62\% \sim 27.1$.
So the answer is 27.6 +/- 27.1.

7. Hellmuth vs. Farha.

P(Hellmuth makes a flush)

$$= \frac{C(11,5) + C(11,4) * 37 + C(11,3) * C(37,2)}{C(48,5)} = 7.16\%.$$

P(Farha makes a flush)

$$= \frac{2 * (C(12,5) + C(12,4) * 36)}{C(48,5)} = 2.17\%.$$