Stat 100a: Introduction to Probability.

Outline for the day:

1. Hand in HW2.
2. Exponential, again.
3. Pareto random variables.
4. E(cards until 2^{nd} king).
5. Covariance and correlation.
6. Hellmuth and Farha.
7. Review list.
8. Practice problems.

Don't forget to bring a calculator, and any books or notes, to the exam this Wed.

♠ ♣ ♥ ♦

1. **HAND IN HW2.**

2. **Exponential distribution, ch 6.4.**

Useful for modeling waiting times til something happens (like the geometric).

pdf of an exponential random variable is $f(y) = \lambda \exp(-\lambda y)$, for $y \geq 0$, and $f(y) = 0$ otherwise.
The cdf is $F(y) = 1 - \exp(-\lambda y)$, for $y \geq 0$.
If X is exponential with parameter λ, then $E(X) = SD(X) = 1/\lambda$

If the total numbers of events in any disjoint time spans are independent, then these totals are Poisson random variables. If in addition the events are occurring at a constant rate λ, then the times between events, or *interevent times*, are exponential random variables with mean $1/\lambda$.

Example. Suppose you play 20 hands an hour, with each hand lasting exactly 3 minutes, and let X be the time in hours until the end of the first hand in which you are dealt pocket aces. Use the exponential distribution to approximate $P(X \leq 2)$ and compare with the exact solution using the geometric distribution.

Answer. Each hand takes 1/20 hours, and the probability of being dealt pocket aces on a particular hand is 1/221, so the rate λ = 1 in 221 hands = 1/(221/20) hours ~ 0.0905 per hour.
Using the exponential model, $P(X \le 2\ hours) = 1 - exp(-2\lambda) \sim 16.556\%$.
This is an approximation, however, since by assumption X is not continuous but must be an integer multiple of 3 minutes.
Let Y = the number of hands you play until you are dealt pocket aces. Using the geometric distribution, $P(X \le 2\ hours) = P(Y \le 40\ hands)$
$= 1 - (220/221)^{40} \sim 16.590\%$.

The survivor function for an exponential random variable is particularly simple: $P(X > c) = \int_c^\infty f(y)dy = \int_c^\infty \lambda\ exp(-\lambda\ y)dy = -exp(-\lambda\ y)]_c^\infty = exp(-\lambda\ c)$.

Like geometric random variables, exponential random variables have the *memorylessness* property: if X is exponential, then for any non-negative values a and b, $P(X > a+b \mid X > a) = P(X > b)$. *(See p115)*.
Thus, with an exponential (or geometric) random variable, if after a certain time you still have not observed the event you are waiting for, then the distribution of the *future*, additional waiting time until you observe the event is the same as the distribution of the *unconditional* time to observe the event to begin with.

3. 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*.

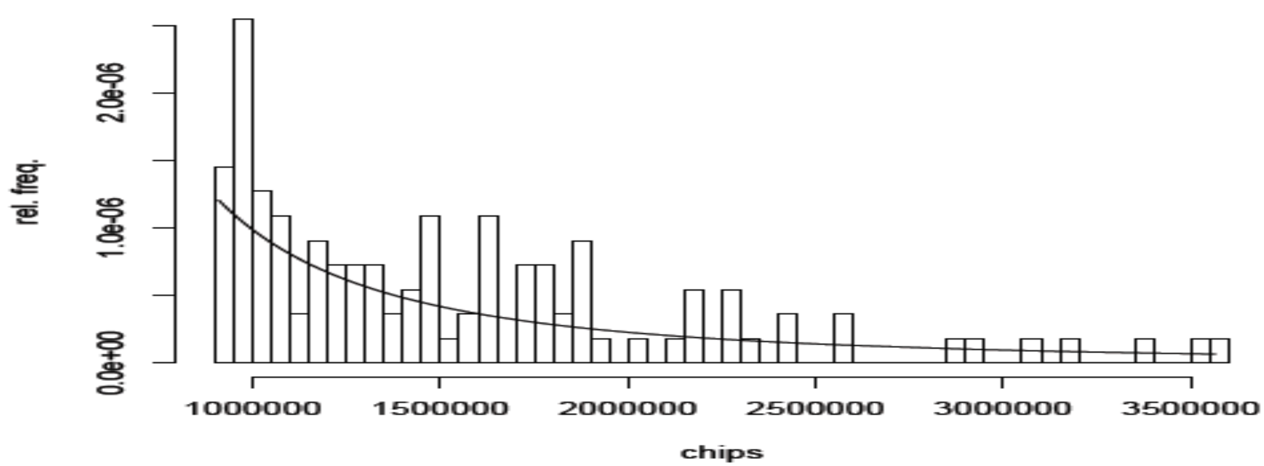

Figure 6.6.1: Relative frequency histogram of the chip counts of the leading 110 players in the 2010 WSOP main event after day 5. The curve is the Pareto density with $a = 900{,}000$ and $b = 1.11$.

4. E(cards til 2nd King).

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

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

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

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

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

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

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

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. Covariance and correlation, p127.

For any random variables X and Y,

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

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

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

cov(X,Y) = E(XY) – E(X)E(Y) is called the *covariance* between X and Y,

cor(X,Y) = 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 cov(X,Y) = 0, and var(X+Y) = var(X) + var(Y).

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

cov(aX + b,Y) = E[(aX+b)Y] – E(aX+b)E(Y)

= aE(XY) + bE(Y) – [aE(X)E(Y) + bE(Y)] = a 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 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. cor(X,Y) = 0.

But X and Y are not independent.

P(X=2 and Y=14) = 0, but P(X=2)P(Y=14) = (1/13)(1/26).

6. Hellmuth and 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)

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

7. Review list.

1) Basic principles of counting.
2) Axioms of probability, and addition rule.
3) Permutations & combinations.
4) Conditional probability.
5) Independence.
6) Multiplication rules. $P(AB) = P(A)\ P(B|A)$ [$= P(A)P(B)$ if ind.]
7) Odds ratios.
8) Random variables (RVs).
9) Discrete RVs, and probability mass function (pmf).
10) Expected value.
11) Pot odds calculations.
12) Luck, skill, and deal-making.
13) Variance and SD.
14) Bernoulli RV. [0-1. $\mu = p$, $s = \sqrt{(pq)}$.]
15) Binomial RV. [# of successes, out of n tries. $\mu = np$, $s = \sqrt{(npq)}$.]
16) Geometric RV. [# of tries til 1st success. $\mu = 1/p$, $s = (\sqrt{q}) / p$.]
17) Negative binomial RV. [# of tries til rth success. $\mu = r/p$, $s = (\sqrt{rq}) / p$.]
18) Poisson RV [# of successes in some time interval. [$\mu = l$, $s = \sqrt{l}$.]
19) E(X+Y), V(X+Y) (ch. 7.1).
20) Bayes's rule (ch. 3.4).
20) Continuous RVs
21) Probability density function (pdf)
22) Uniform RV
23) Normal RV
24) Exponential RV
25) Moment generating functions
26) Markov and Chebyshev inequalities
27) Pareto random variables.
28) Correlation, covariance.

Basically, we've done all of 1-7.1 in the book, except 6.7.

8. Some example problems.

a. Find the probability you are dealt a suited king.

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

b. 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.

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.

P(You have AK | you have exactly one ace)?

= P(You have AK and exactly one ace) / P(exactly one ace)

= P(AK) / P(exactly one ace)

= (16/C(52,2)) ÷ (4x48/C(52,2))

= 4/48 = 8.33%.

P(You have AK | you have at least one ace)?

= P(You have AK and at least one ace) / P(at least one ace)

= P(AK) / P(at least one ace)

= (16/C(52,2)) ÷ (((4x48 + C(4,2))/C(52,2)) ~ 8.08%.

P(You have AK | your FIRST card is an ace)?

= 4/51 = 7.84%.