Stat 100a, Introduction to Probability.

Outline for the day:

0. Exam 2.
1. Survivor functions.
2. Pareto.
3. Bivariate normal in *R*.
4. A quick fact about normals.
5. Conditional expectation.
6. Law of Large Numbers.
7. Bivariate and marginal density.

We are skipping 6.7 and the bulk of 6.3 about optimal play with uniform hands.
Read through chapter 7.2.
No class Mon Sep 6 Labor Day!
I will have a short, optional review session and tournament with your *R* functions on Tue Sep7 from 10am to 1030am and will post it on the course website.
HW3 is on the course website and is due Tue Sep7 9am.
For the final exam Sep8, you will need to be on zoom and have your cameras on during the exam!

On your exams, the grading scale is the usual,

96.7-100 = A+,

93.3-96.7 = A,

90-93.3 = A-,

86.7-90 = B+,

83.3-86.7 = B, etc.

I keep a record of your score, not the letter grade.

I do reward improvement on the exams. I will not completely ignore your first midterm, but I do reward improvement.

The exams are cumulative.

1. Survivor functions.

Recall the cdf $F(b) = P(X \leq b)$.
The survivor function is $S(b) = P(X > b) = 1 - F(b)$.
Some random variables have really simple survivor functions and it can be convenient to work with them.
If X is geometric, then $S(b) = P(X > b) = q^b$, for $b = 0,1,2,....$
For instance, let b=2. $X > 2$ means the 1st two were misses,
i.e. $P(X>2) = q^2$.
For exponential X, $F(b) = 1 - exp(-\lambda b)$, so $S(b) = exp(-\lambda b)$.

An interesting fact is that, if X takes only values in $\{0,1,2,3,...\}$,
then $E(X) = S(0) + S(1) + S(2) + ...$
Proof.

$$
\begin{aligned}
S(0) &= P(X=1) + P(X=2) + P(X=3) + P(X=4) + ... \\
S(1) &= \phantom{P(X=1) + {}} P(X=2) + P(X=3) + P(X=4) + ... \\
S(2) &= \phantom{P(X=1) + P(X=2) + {}} P(X=3) + P(X=4) + ... \\
S(3) &= \phantom{P(X=1) + P(X=2) + P(X=3) + {}} P(X=4) + ...
\end{aligned}
$$

Add these up and you get
$0\,P(X=0) + 1P(X=1) + 2P(X=2) + 3P(X=3) + 4P(X=4) + ...$
$= \sum kP(X=k) = E(X)$.

2. Pareto distribution, ch 6.6.

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*.

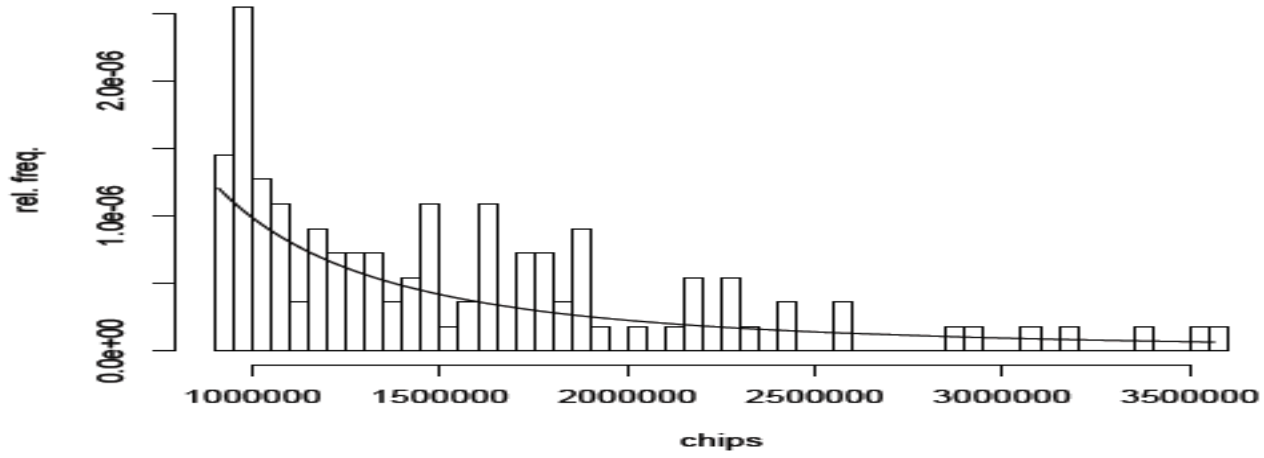


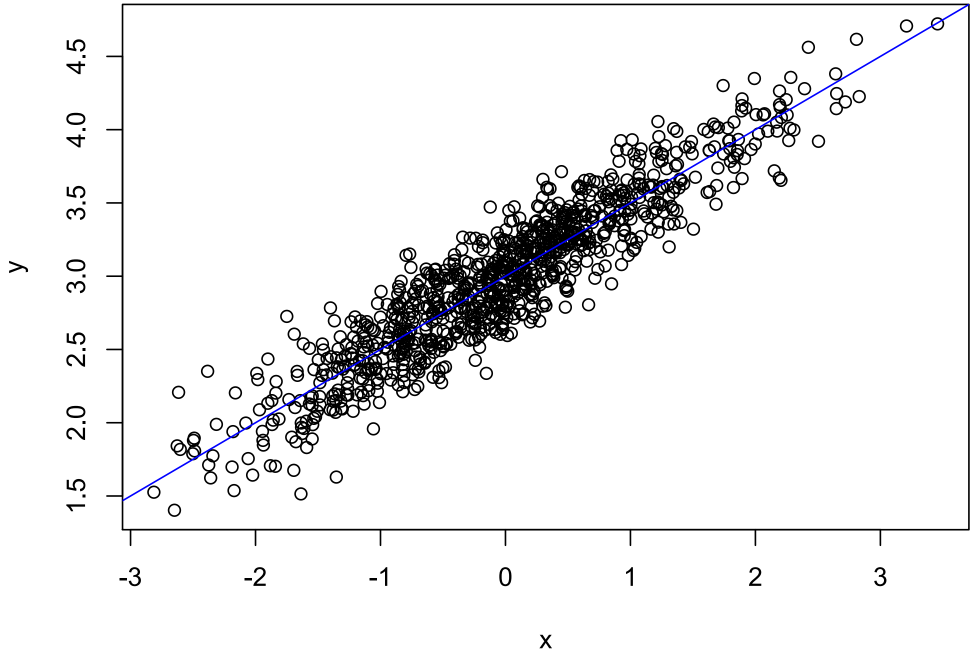
Figure 6.6.1: Relative frequency histogram of the chip counts of the leading 110 players in the 2010 WSOP main event after day 5. The curve is the Pareto density with $a = 900{,}000$ and $b = 1.11$.

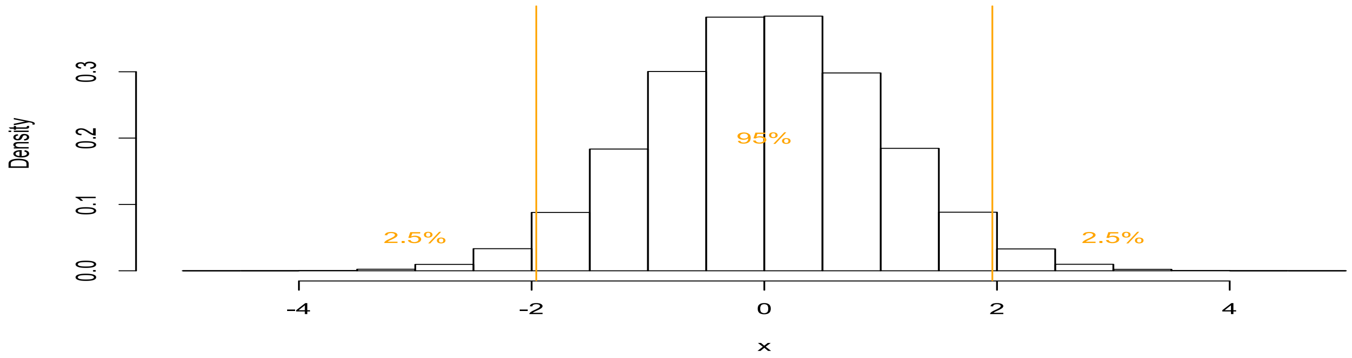
3. Bivariate normal.

For example, let X = N(0,1). Let ε = N(0, 0.2^2) and independent of X. Let Y = 3 + 0.5 X + ε.

In R,

```
x = rnorm(1000,mean=0,sd=1)
eps = rnorm(1000,mean=0,sd=.2)
y = 3 + .5*x+eps
plot(x,y)
cor(x,y) # 0.9282692.
```





4. A quick fact about normals.

If X and Y are independent and both are normal, then X+Y is normal, and so are –X and –Y.

5. 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%.

6. Law of Large Numbers (LLN) and the Fundamental Theorem of Poker.

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.

7. Bivariate and marginal density.

Suppose X and Y are random variables.

If X and Y are discrete, we can define the joint pmf $f(x,y) = P(X = x \text{ and } Y = y)$.

Suppose X and Y are continuous for the rest of this page.

Define the bivariate or joint pdf $f(x,y)$ as a function with the properties that $f(x,y) \geq 0$, and for any a,b,c,d,

$$P(a \leq X \leq b \text{ and } c \leq Y \leq d) = \int_a^b \int_c^d f(x,y)\, dy\, dx.$$

The integral $\int_{-\infty}^{\infty} f(x,y)\, dy = f(x)$, the pdf of X, and this function $f(x)$ is sometimes called the *marginal* density of X. Similarly $\int_{-\infty}^{\infty} f(x,y)\, dx$ is the marginal pdf of Y.

$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-\infty}^{\infty} x\, [\int_{-\infty}^{\infty} f(x,y)\, dy]\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f(x,y)\, dy\, dx$.

Just as $P(A|B) = P(AB)/P(B)$, $f(x|y) = f(x,y)/f(y)$.

X and Y are independent iff. $f(x,y) = f_x(x)f_y(y)$.

Now $E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x,y)\, dy\, dx$. This can be useful to find $\text{cov}(X,Y) = E(XY) - E(X)E(Y)$.

What is $E(X^2Y+e^Y)$? It $= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x^2y+e^y)\, f(x,y)\, dy\, dx$.

Bivariate and marginal density.

Suppose the joint density of X and Y is $f(x,y) = a \exp(x+y)$, for X and Y in (0,1) x (0,1). What is a? What is the marginal density of Y? What type of distribution does X have conditional on Y? What is E(X|Y)? What is the mean of X when Y = .2? Are X and Y independent?

$\iint a \exp(x+y)\, dxdy = 1 = a\iint \exp(x)\exp(y)dxdy = a\int_0^1 \exp(x)\, dx \int_0^1 \exp(y)\, dy = a(e-1)^2$,
so $a = (e-1)^{-2}$.

The marginal density of Y is $f(y) = \int_0^1 a \exp(x+y)\, dx = a \exp(y) \int_0^1 \exp(x)dx = a \exp(y)(e-1) = \exp(y)/(e-1)$.

Conditional on Y, the density of X is $f(x|y) = f(x,y)/f(y) = a \exp(x+y)(e-1)/\exp(y) = \exp(x)/(e-1)$. So X|Y is like an exponential(1) random variable restricted to (0,1).

$E(X|Y) = \int_0^1 x \exp(x)/(e-1)\, dx = 1/(e-1)\, [x \exp(x) - \int \exp(x)dx] = 1/(e-1)\, [x \exp(x) – \exp(x)]_0^1 = 1/(e-1)\, [e – e – 0 + 1] = 1/(e-1)$.
When Y = .2, $E(X|Y) = 1/(e-1)$.

$f(y) = \exp(y)/(e-1)$ and similarly $f(x) = \exp(x)/(e-1)$,
so $f(x)f(y) = \exp(x+y)/(e-1)^2 = f(x,y)$. Therefore, X and Y are independent.