

Stat 100a: Introduction to Probability.

Outline for the day:

1. Hand back midterm 1, and other announcements.
2. Harman and Negreanu.
3. Continuous random variables.
4. Exponential random variables.
5. Uniform random variables.
6. Normal random variables.
7. Functions of independent random variables.
8. Moment generating functions of some random variables.
9. Survivor functions.
10. Deal making.
11. $E(\text{cards til } 2^{\text{nd}} \text{ king})$.

Hw2 and the outline at the bottom on the syllabus on the course website has been slightly revised.

Hw2 is 4.6, 4.8, 4.16, 5.2. I removed 5.6.

Read through chapter 7.2 for next week.



1. Midterm 1.

BE QUIET UNTIL I AM DONE HANDING OUT ALL EXAMS.

In computing your final grade at the end of the quarter, I use the standard grading scale, where 90-100 = A range

80s = B range, 70s = C range, 60s = D range, below 60 = F.

Hw2 and the outline at the bottom on the syllabus on the course website has been slightly revised. Hw2 is 4.6, 4.8, 4.16, 5.2. I removed 5.6. Midterm 2 is still Aug 24 in class, and the project is still due to me by email on Sun Aug 28 8pm.

The midterm is cumulative and is on chapters 1-6 and 7.1, but not 6.6 and 6.7.

All emails are on the course website in emails.txt.

2. Harman / Negreanu, and running it twice.

Harman has $10\spadesuit 7\spadesuit$. Negreanu has $K\heartsuit Q\heartsuit$. The flop is $10\diamondsuit 7\clubsuit K\diamondsuit$.

Harman's all-in. \$156,100 pot. $P(\text{Negreanu wins}) = 28.69\%$. $P(\text{Harman wins}) = 71.31\%$.

Let X = amount Harman has after the hand.

If they run it once, $E(X) = \$0 \times 29\% + \$156,100 \times 71.31\% = \mathbf{\$111,314.90}$.

If they run it twice, what is $E(X)$?

There's some probability p_1 that Harman wins both times $\implies X = \$156,100$.

There's some probability p_2 that they each win one $\implies X = \$78,050$.

There's some probability p_3 that Negreanu wins both $\implies X = \$0$.

$E(X) = \$156,100 \times p_1 + \$78,050 \times p_2 + \$0 \times p_3$.

If the different runs were *independent*, then $p_1 = P(\text{Harman wins 1st run \& 2nd run})$

would $= P(\text{Harman wins 1st run}) \times P(\text{Harman wins 2nd run}) = 71.31\% \times 71.31\% \sim 50.85\%$.

But, they're not quite independent! Very hard to compute p_1 and p_2 .

However, you don't need p_1 and p_2 !

X = the amount Harman gets from the 1st run + amount she gets from 2nd run, so

$E(X) = E(\text{amount Harman gets from 1st run}) + E(\text{amount she gets from 2nd run})$

$= \$78,050 \times P(\text{Harman wins 1st run}) + \$0 \times P(\text{Harman loses first run})$

$+ \$78,050 \times P(\text{Harman wins 2nd run}) + \$0 \times P(\text{Harman loses 2nd run})$

$= \$78,050 \times 71.31\% + \$0 \times 28.69\% + \$78,050 \times 71.31\% + \$0 \times 28.69\% = \mathbf{\$111,314.90}$.

HAND RECAP Harman $10\spadesuit 7\spadesuit$ Negreanu $K\heartsuit Q\heartsuit$ The flop is $10\diamondsuit 7\clubsuit K\diamondsuit$.

Harman's all-in. \$156,100 pot. $P(\text{Negreanu wins}) = 28.69\%$. $P(\text{Harman wins}) = 71.31\%$.

The standard deviation (SD) changes a lot! **Say they run it once.** (see p127.)

$$V(X) = E(X^2) - \mu^2.$$

$$\mu = \$111,314.9, \text{ so } \mu^2 \sim \$12.3 \text{ billion.}$$

$$E(X^2) = (\$156,100^2)(71.31\%) + (0^2)(28.69\%) = \$17.3 \text{ billion.}$$

$$V(X) = \$17.3 \text{ billion} - \$12.3 \text{ bill.} = \$5.09 \text{ billion. SD } \sigma = \sqrt{\$5.09 \text{ billion}} \sim \$71,400.$$

So if they run it once, Harman expects to get back about \$111,314.9 +/- **\$71,400.**

If they run it twice? Hard to compute, but approximately, if each run were

independent, then $V(X_1 + X_2) = V(X_1) + V(X_2)$,

so if X_1 = amount she gets back on 1st run, and X_2 = amount she gets from 2nd run,

then $V(X_1 + X_2) \sim V(X_1) + V(X_2) \sim \$1.25 \text{ billion} + \$1.25 \text{ billion} = \2.5 billion ,

The standard deviation $\sigma = \sqrt{\$2.5 \text{ billion}} \sim \$50,000$.

So if they run it twice, Harman expects to get back about \$111,314.9 +/- **\$50,000.**

3. Continuous random variables and their densities, p103-107.

Density (or pdf = Probability Density Function) $f(y)$:

$$\int_B f(y) dy = P(X \text{ in } B).$$

By the fundamental theorem of calculus, $f(y) = F'(y)$, where $F(y)$ is the cumulative distribution function.

Expected value, $\mu = E(X) = \int y f(y) dy$. (= $\sum y P(y)$ for discrete X .)

Variance, $\sigma^2 = V(X) = E(X^2) - \mu^2$.

$SD(X) = \sqrt{V(X)}$.

For examples of pdfs, see p104, 106, and 107.

4. 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.

5. Uniform Random Variables and R.

Continuous random variables are often characterized by their

probability density functions (pdf, or *density*): a function $f(x)$

such that $P\{X \text{ is in } B\} = \int_B f(x) dx$.

Uniform: $f(x) = c$, for x in (a, b) .

$= 0$, for all other x .

[Note: c must $= 1/(b-a)$, so that $\int_a^b f(x) dx = P\{X \text{ is in } (a,b)\} = 1$.]

Uniform $(0,1)$. See p107-109.

$f(y) = 1$, for y in $(0,1)$. $\mu = 0.5$. $\sigma \sim 0.29$.

$P(X \text{ is between } 0.4 \text{ and } 0.6) = \int_{.4}^{.6} f(y) dy = \int_{.4}^{.6} 1 dy = 0.2$.

In R, `runif(1,min=a,max=b)` produces a pseudo-random uniform.

Uniform example.

For a continuous random variable X ,

The pdf $f(y)$ is a function where $\int_a^b f(y)dy = P\{X \text{ is in } (a,b)\}$,

$$E(X) = \mu = \int_{-\infty}^{\infty} y f(y)dy,$$

$$\text{and } \sigma^2 = \text{Var}(X) = E(X^2) - \mu^2. \quad \text{sd}(X) = \sigma.$$

For example, suppose X and Y are independent uniform random variables on $(0,1)$, and $Z = \min(X,Y)$. **a)** Find the pdf of Z . **b)** Find $E(Z)$. **c)** Find $SD(Z)$.

a. For c in $(0,1)$, $P(Z > c) = P(X > c \text{ \& } Y > c) = P(X > c) P(Y > c) = (1-c)^2 = 1 - 2c + c^2$.
So, $P(Z \leq c) = 1 - (1 - 2c + c^2) = 2c - c^2$.

Thus, $\int_0^c f(c)dc = 2c - c^2$. So $f(c)$ = the derivative of $2c - c^2 = 2 - 2c$, for c in $(0,1)$.
Obviously, $f(c) = 0$ for all other c .

$$\begin{aligned} \text{b. } E(Z) &= \int_{-\infty}^{\infty} y f(y)dy = \int_0^1 c (2-2c) dc = \int_0^1 2c - 2c^2 dc = c^2 - 2c^3/3 \Big|_{c=0}^1 \\ &= 1 - 2/3 - (0 - 0) = 1/3. \end{aligned}$$

$$\begin{aligned} \text{c. } E(Z^2) &= \int_{-\infty}^{\infty} y^2 f(y)dy = \int_0^1 c^2 (2-2c) dc = \int_0^1 2c^2 - 2c^3 dc = 2c^3/3 - 2c^4/4 \Big|_{c=0}^1 \\ &= 2/3 - 1/2 - (0 - 0) = 1/6. \end{aligned}$$

$$\text{So, } \sigma^2 = \text{Var}(Z) = E(Z^2) - [E(Z)]^2 = 1/6 - (1/3)^2 = 1/18.$$

$$SD(Z) = \sigma = \sqrt{(1/18)} \sim 0.2357.$$

6. Normal random variables.

So far we have seen two continuous random variables, the uniform and the exponential.

Normal. pp 115-117. mean = μ , SD = σ , $f(y) = 1/\sqrt{(2\pi\sigma^2)} e^{-(y-\mu)^2/2\sigma^2}$.

Symmetric around μ ,

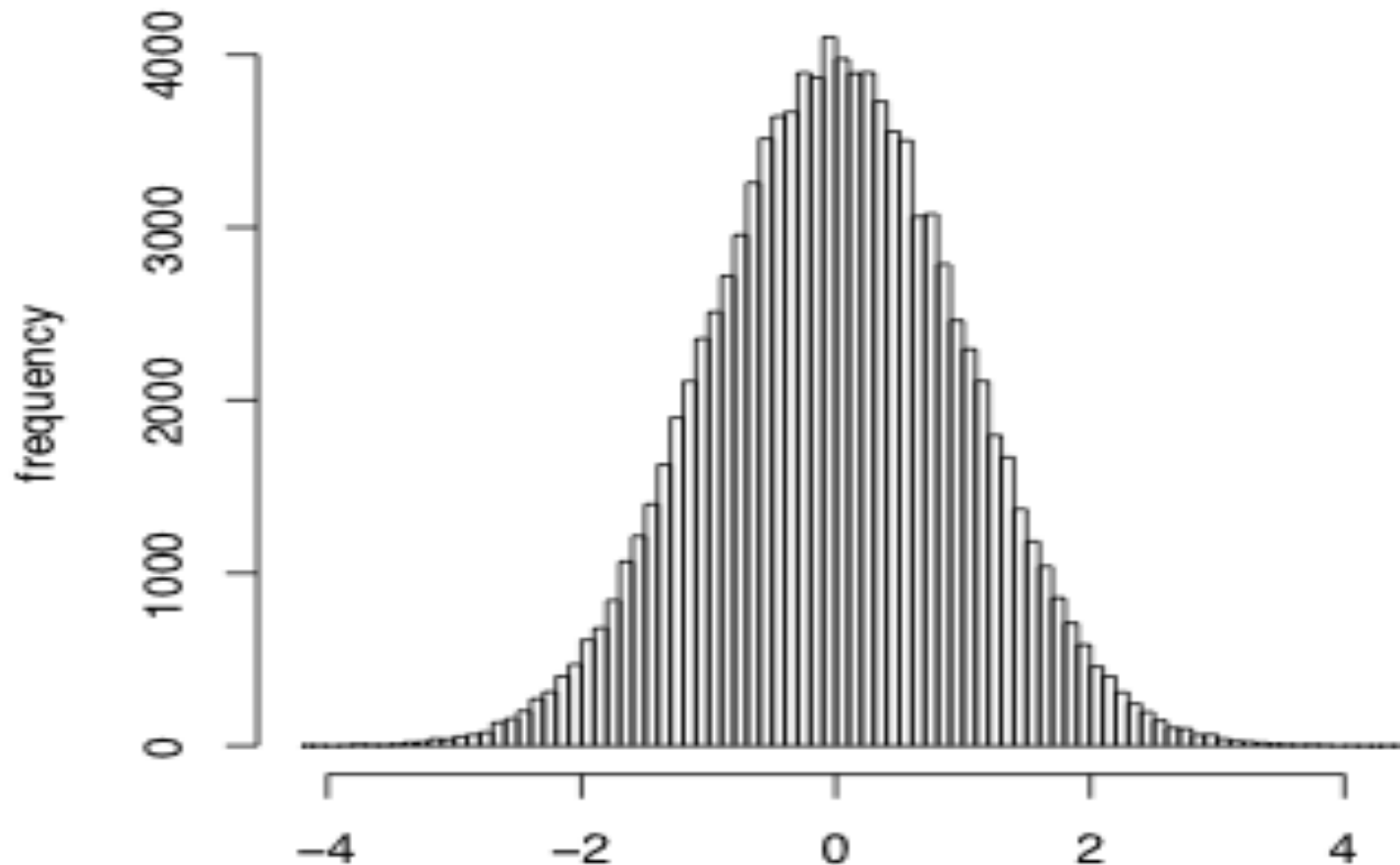
50% of the values are within 0.674 SDs of μ ,

68.27% of the values are within 1 SD of μ , and

95% are within 1.96 SDs of μ .

* Standard Normal. Normal with $\mu = 0$, $\sigma = 1$. See pp 117-118.

Standard normal density:
68.27% between -1.0 and 1.0
95% between -1.96 and 1.96



7. Functions of independent random variables.

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 problem 5.4.

8. Moment generating functions of some random variables.

Bernoulli(p). $\phi_X(t) = pe^t + q$.

Binomial(n,p). $\phi_X(t) = (pe^t + q)^n$. p94.

Geometric(p). $\phi_X(t) = pe^t/(1 - qe^t)$.

Neg. binomial (r,p). $\phi_X(t) = [pe^t/(1 - qe^t)]^r$. p97.

Poisson(λ). $\phi_X(t) = e^{\{\lambda e^t - \lambda\}}$. p100.

Uniform (a,b). $\phi_X(t) = (e^{tb} - e^{ta})/[t(b-a)]$. p108.

Exponential (λ). $\phi_X(t) = \lambda/(\lambda - t)$. p123.

Normal. $\phi_X(t) = e^{\{t\mu + t^2\sigma^2/2\}}$.

Note t is missing in neg. binomial one on p97.

9. Survivor functions. p96 and 115.

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 $b = 0, 1, 2, \dots$

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 on only values in $\{0, 1, 2, 3, \dots\}$,
then $E(X) = S(0) + S(1) + S(2) + \dots$

Proof. See p96.

$$S(0) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + \dots$$

$$S(1) = P(X=2) + P(X=3) + P(X=4) + \dots$$

$$S(2) = P(X=3) + P(X=4) + \dots$$

$$S(3) = P(X=4) + \dots$$

Add these up and you get

$$0 P(X=0) + 1P(X=1) + 2P(X=2) + 3P(X=3) + 4P(X=4) + \dots$$

$$= \sum kP(X=k) = E(X).$$

10. Deal making. (Expected value, game theory)

Game-theory: For a symmetric-game tournament, the probability of winning is approx. optimized by the *myopic rule* (in each hand, maximize your expected number of chips),
and

$P(\text{you win}) = \text{your proportion of chips}$ (Theorems 7.6.6 and 7.6.7 on pp 151-152).

For a *fair* deal, the amount you win = the *expected value* of the amount you will win.
See p61.

For instance, suppose a tournament is winner-take-all, for \$8600.

With 6 players left, you have 1/4 of the chips left.

An *EVEN SPLIT* would give you $\$8600 \div 6 = \mathbf{\$1433}$.

A *PROPORTIONAL SPLIT* would give you $\$8600 \times (\text{your fraction of chips})$
 $= \$8600 \times (1/4) = \mathbf{\$2150}$.

A *FAIR DEAL* would give you the *expected value* of the amount you will win
 $= \$8600 \times P(\text{you get 1st place}) = \mathbf{\$2150}$.

But suppose the tournament is not winner-take-all, but pays
\$3800 for 1st, \$2000 for 2nd, \$1200 for 3rd, \$700 for 4th, \$500 for 5th, \$400 for 6th.

Then a *FAIR DEAL* would give you

$\$3800 \times P(1\text{st place}) + \$2000 \times P(2\text{nd}) + \$1200 \times P(3\text{rd}) + \$700 \times P(4\text{th}) + \$500 \times P(5\text{th}) + \$400 \times P(6\text{th})$.

Hard to determine these probabilities. But, $P(1\text{st}) = 25\%$, and you might roughly
estimate the others as $P(2\text{nd}) \sim 20\%$, $P(3\text{rd}) \sim 20\%$, $P(4\text{th}) \sim 15\%$, $P(5\text{th}) \sim 10\%$, $P(6\text{th})$
 $\sim 10\%$, and get

$\$3800 \times 25\% + \$2000 \times 25\% + \$1200 \times 20\% + \$700 \times 15\% + \$500 \times 10\% + \$400 \times 5\% = \mathbf{\$1865}$.

If you have 40% of the chips in play, then:

EVEN SPLIT = **\$1433**.

PROPORTIONAL SPLIT = **\$3440**.

FAIR DEAL ~ **\$2500!**

Another example. Before the Wasicka/Binger/Gold hand,

Gold had 60M, Wasicka 18M, Binger 11M.

Payouts: 1st place \$12M, 2nd place \$6.1M, 3rd place \$4.1M.

Proportional split: of the total prize pool left, you get your proportion of chips in play.

e.g. \$22.2M left, so Gold gets $60\text{M}/(60\text{M}+18\text{M}+11\text{M}) \times \$22.2\text{M} \sim \$15.0\text{M}$.

A fair deal would give you

$P(\text{you get 1st place}) \times \$12\text{M} + P(\text{you get 2nd place}) \times \$6.1\text{M} + P(\text{3rd pl.}) \times \$4.1\text{M} .$

*Even split: Gold \$7.4M, Wasicka \$7.4M, Binger \$7.4M.

*Proportional split: Gold \$15.0M, Wasicka \$4.5M, Binger \$2.7M.

*Fair split: Gold \$10M, Wasicka \$6.5M, Binger \$5.7M.

*End result: Gold \$12M, Wasicka \$6.1M, Binger \$4.1M.

11. E(cards til 2nd king).

Deal the cards face up, without reshuffling.

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

What is $E(Z)$?

I'll answer this next time. The solution uses the fact

$$E(X+Y+Z + \dots) = E(X) + E(Y) + E(Z) + \dots$$