Stat 100a: Introduction to Probability.

Outline for the day:

1. Moment generating function of uniform.
2. Correlation.
3. Bivariate normal examples.
4. Conditional expectation.
5. LLN.
6. Review list.

We are skipping 6.7 and the bulk of 6.3 about optimal play with uniform hands.

1. Moment generating function of a uniform random variable.

If X is uniform(a,b), then it has density f(x) = 1/(b-a) between a and b, and f(x) = 0 for all other x.

$\phi_X(t) = E(e^{tX})$

$= \int_a^b e^{tx} f(x)\, dx$

$= \int_a^b e^{tx}\, 1/(b-a)\, dx$

$= 1/(b-a) \int_a^b e^{tx}\, dx$

$= 1/(b-a)\, e^{tx}/t]_a^b\, dx$

$= (e^{tb} - e^{ta})/[t(b-a)].$

2. Correlation and covariance.

For any random variables X and Y, recall

var(X+Y) = var(X) + var(Y) + 2cov(X,Y).

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

cor(X,Y) = cov(X,Y) / [SD(X) SD(Y)] is the *correlation* bet. X and Y.

For any real numbers a and b, E(aX + b) = aE(X) + b, and

cov(aX + b,Y) = a cov(X,Y).

var(aX+b) = cov(aX+b, aX+b) = a^2var(X).

No such simple statement is true for correlation.

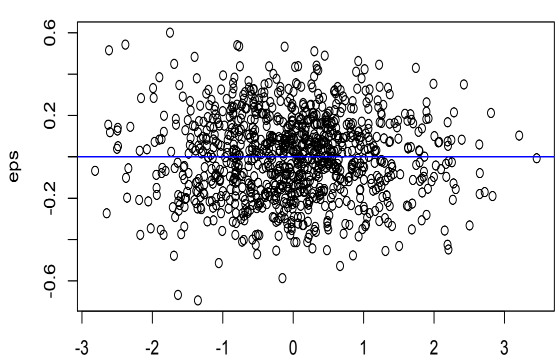
If ρ = cor(X,Y), we always have $-1 \le \rho \le 1$.

$\rho = -1$ iff. the points (X,Y) all fall exactly on a line sloping downward, and

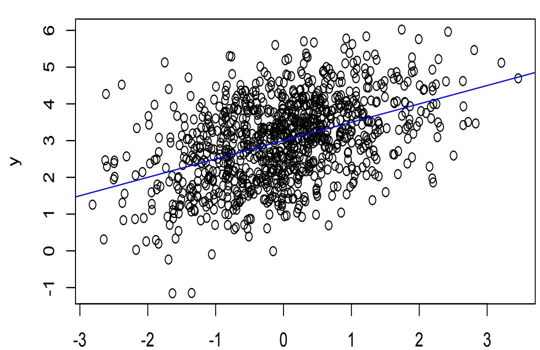
$\rho = 1$ iff. the points (X,Y) all fall exactly on a line sloping upward.

$\rho = 0$ means the best fitting line to (X,Y) is horizontal.

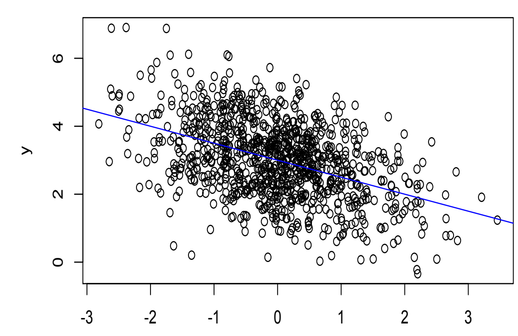
$\rho = 0$



$\rho = 0.44$



$\rho = -0.44$.



Bivariate normal.

If (X,Y) are bivariate normal with E(X) = 100, var(X) = 25, E(Y) = 200, var(Y) = 49, $\rho = 0.8$,

What is the distribution of Y given X = 105? What is P(Y > 213.83 | X = 105)?

Given X = 105, Y is normal. Write $Y = \beta_1 + \beta_2 X + \varepsilon$ where ε is normal with mean 0, ind. of X.

Recall $\beta_2 = \rho\, \sigma_y/\sigma_x = 0.8\ (7/5) = 1.12$.

So $Y = \beta_1 + 1.12\ X + \varepsilon$.

To get β_1, note $200 = E(Y) = \beta_1 + 1.12\ E(X) + E(\varepsilon) = \beta_1 + 1.12\ (100)$. So $200 = \beta_1 + 112$. $\beta_1 = 88$.

So $Y = 88 + 1.12\ X + \varepsilon$, where ε is normal with mean 0 and ind. of X.

What is var(ε)?
$49 = \mathrm{var}(Y) = \mathrm{var}(88 + 1.12\ X + \varepsilon) = 1.12^2\ \mathrm{var}(X) + \mathrm{var}(\varepsilon) + 2(1.12)\ \mathrm{cov}(X,\varepsilon)$
$= 1.12^2\ (25) + \mathrm{var}(\varepsilon) + 0$. So $\mathrm{var}(\varepsilon) = 49 - 1.12^2\ (25) = 17.64$ and $\mathrm{sd}(\varepsilon) = \sqrt{17.64} = 4.2$.

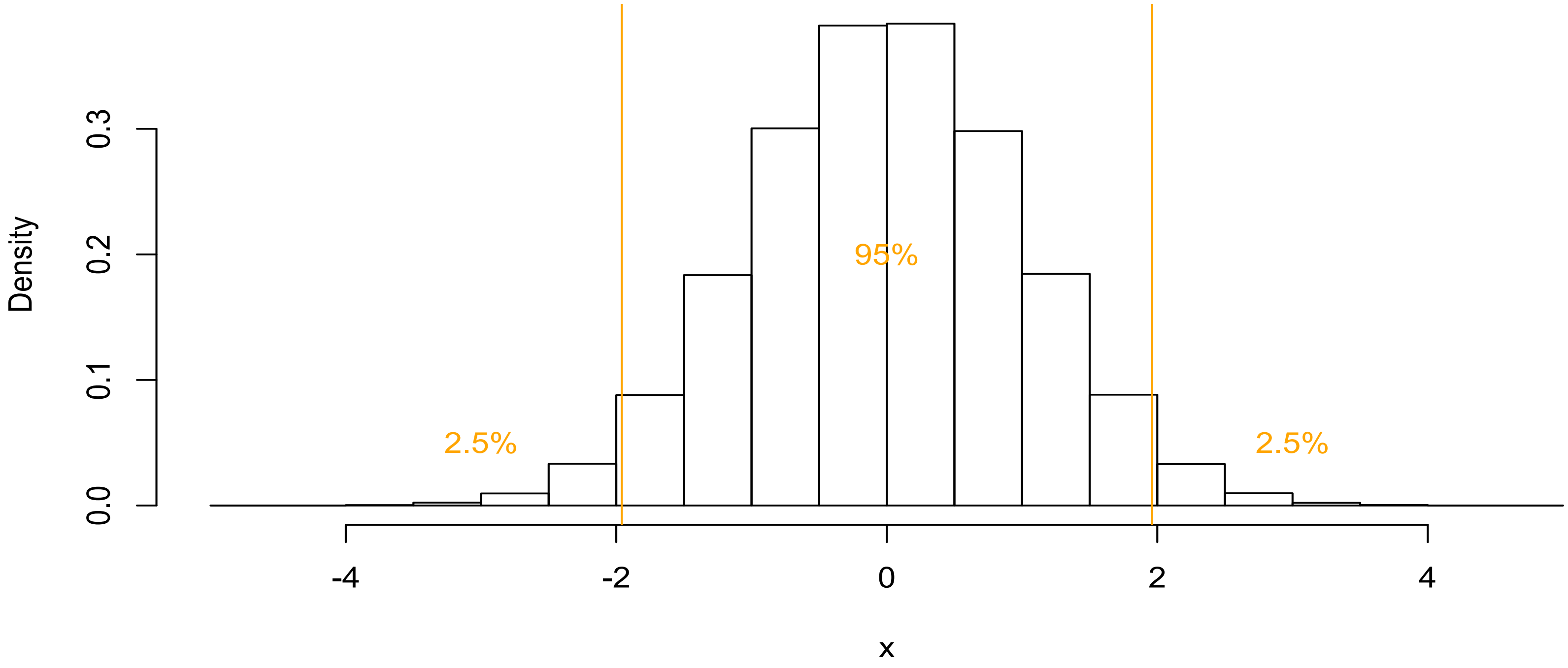
So $Y = 88 + 1.12\ X + \varepsilon$, where ε is $N(0, 4.2^2)$ and ind. of X.

Given X = 105, $Y = 88 + 1.12(105) + \varepsilon = 205.6 + \varepsilon$, so $Y|X{=}105 \sim N(205.6, 4.2^2)$.
Now how many sds above the mean is 213.83? $(213.83 - 205.6)/4.2 = 1.96$,
so P(Y>213.83 | X=105) = P(normal is > 1.96 sds above its mean) = 2.5%.

Bivariate normal.

Now how many sds above the mean is 213.83? (213.83 – 205.6)/4.2 = 1.96,

so P(Y>213.83 | X=105) = P(normal is > 1.96 sds above its mean) = 2.5%.



More example problems.

X is a continuous random variable with cdf $F(y) = 1 - y^{-1}$, for y in $(1,\infty)$, and $F(y) = 0$ otherwise.
a. What is the pdf of X?
b. What is f(1)?
c. What is E(X)?

a. $f(y) = F'(y) = d/dy\,(1 - y^{-1}) = y^{-2}$, for y in $(1,\infty)$, and $f(y) = 0$ otherwise.

To be a pdf, f(y) must be nonnegative for all y and integrate to 1.
$f(y) \geq 0$ for all y, and $\int_{-\infty}^{\infty} f(y)dy = \int_1^{\infty} y^{-2}\,dy = -y^{-1}]_1^{\infty} = 0 + 1 = 1$. So, f is indeed a pdf.

b. $f(1) = 1^{-2} = 1$.

c. $E(X) = \int_{-\infty}^{\infty} y\, f(y)dy = \int_1^{\infty} y\, y^{-2}\, dy = \int_1^{\infty} y^{-1}\, dy = \ln(\infty) - \ln(1) = \infty$. This can happen.

X is a continuous random variable with cdf $F(y) = 1 - y^{-2}$, for y in $(1,\infty)$, and $F(y) = 0$ otherwise.
a. What is the pdf of X?
b. What is f(1)? Is this a problem?
c. What is E(X)?
d. What is $P(2 \le X \le 3)$?
e. What is $P(2 < X < 3)$?

a. $f(y) = F'(y) = d/dy\,(1 - y^{-2}) = 2y^{-3}$, for y in $(1,\infty)$, and $f(y) = 0$ otherwise.

To be a pdf, f(y) must be nonnegative for all y and integrate to 1.
$f(y) \ge 0$ for all y, and $\int_{-\infty}^{\infty} f(y)dy = \int_1^{\infty} 2y^{-3}\,dy = -y^{-2}]_1^{\infty} = 0 + 1 = 1$. So, f is indeed a pdf.

b. $f(1) = 2$. This does not mean $P(X=1)$ is 2. It is not a problem.

c. $E(X) = \int_{-\infty}^{\infty} y\, f(y)dy = 2\int_1^{\infty} y\, y^{-3}\,dy = 2\int_1^{\infty} y^{-2}\,dy = -2y^{-1}]_1^{\infty} = 0 + 2 = 2$.

d. $P(2 \le X \le 3) = \int_2^3 f(y)dy = 2\int_2^3 y^{-3}\,dy = -y^{-2}\,]_2^3 = -1/9 + 1/4 \sim 0.139$.

Alternatively, $P(2 \le X \le 3) = F(3) – F(2) = 1 – 1/9 – 1 + 1/4 \sim 0.139$.
e. Same thing.

Suppose X is uniform on (0,1), Y is exponential with E(Y)=2, and X and Y are independent. What is cov(3X+Y, 4X-Y)?

cov(3X+Y, 4X-Y) = 12 cov(X,X) – 3cov(X,Y) + 4cov(Y,X) – cov(Y,Y)
= 12 var(X) – 0 + 0 – var(Y).

For exponential, E(Y) = $1/\lambda$ and var(Y) = $1/\lambda^2$, so $\lambda=1/2$ and var(Y) = 4.
What about var(X)?
$E(X^2) = \int y^2 f(y)dy$
$= \int_0^1 y^2 dy$ because f(y) = 1 for uniform(0,1) for y in (0,1),
$= y^3/3]_0^1$
$= 1/3.$
var(X) = $E(X^2) - \mu^2 = 1/3 – 1/4 = 1/12$.

cov(3X+Y, 4X-Y) = 12 (1/12) – 4
= -3.

Conditional expectation, E(Y | X), ch. 7.2.
Suppose X and Y are discrete.
Then E(Y | X=j) is defined as Σ_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%.

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.

2. 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) Discrete RVs and probability mass function (pmf).
9) Expected value.
10) Pot odds calculations.
11) Luck and skill.
12) Variance and SD.
13) $E(aX+b)$ and $E(X+Y)$.
14) Bayes's rule.
15) Markov and Chebyshev inequalities.
16) Bernoulli, binomial, geometric, Poisson, and negative binomial rvs.
17) Moment generating functions.
18) pdf, cdf, uniform, exponential, normal and Pareto rvs. $F'(y) = f(y)$.
19) Survivor functions.
20) Covariance and correlation.
21) Bivariate normal.
22) Conditional expectation.
23) LLN.

We have basically done all of chapters 1-7.2. Ignore 6.7 and most of 6.3 on optimal play.

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.