Stat 100a, Introduction to Probability.

Outline for the day.

1. Hand in hw2.
2. Bivariate normal.
3. More about covariance and correlation.
4. Example problems.

Bring a PEN or PENCIL and CALCULATOR and any books or notes you want to the exams.

- No class Thu Nov 11 Veteran's Day, and Thu Nov25 Thanksgiving.
- Next midterm is Tue Nov16 in class.
- HW2 is due today, Tue Nov9, and can submitted as pdf via CCLE.
- HW3 is on the course website.
- Please write your ID number at the top of your hw and leave all answers as decimals, not fractions.
- Tue Nov23, lecture will be on zoom, and no office hour that day. Use the zoom link https://ucla.zoom.us/j/91509411456?pwd=aXNUMmhYRElBUzljeXdPMHExSkljZz09
  Meeting ID: 915 0941 1456. Password: 235711

http://www.stat.ucla.edu/~frederic/100A/F21 .

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,

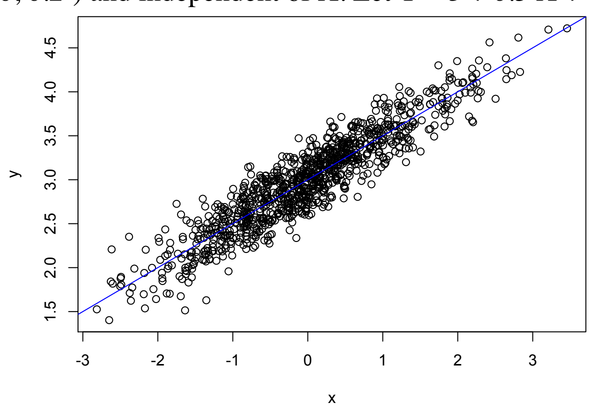
x = rnorm(1000,mean=0,sd=1)

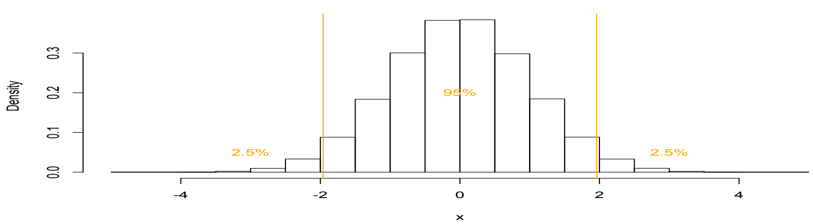
eps = rnorm(1000,mean=0,sd=.2)

y = 3 + .5*x+eps

plot(x,y)

cor(x,y) # 0.9282692.





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 $\beta_2 = \rho\ \sigma_y/\sigma_x = 0.8 \times 7/5 = 1.12$.

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.

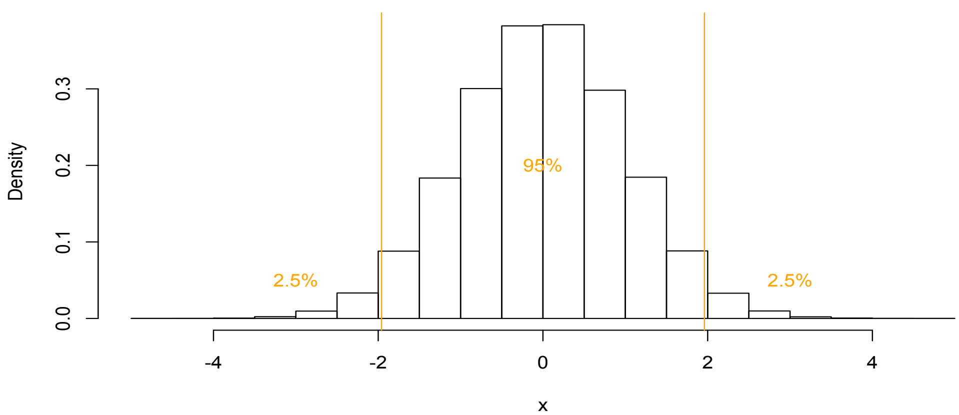
What is var(ε)?
$49 = \text{var}(Y) = \text{var}(88 + 1.12\ X + \varepsilon) = 1.12^2\ \text{var}(X) + \text{var}(\varepsilon) + 2(1.12)\ \text{cov}(X,\varepsilon)$
$= 1.12^2\ (25) + \text{var}(\varepsilon) + 0$. So $\text{var}(\varepsilon) = 49 - 1.12^2\ (25) = 17.64$ and $\text{sd}(\varepsilon) = \sqrt{17.64}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

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%.



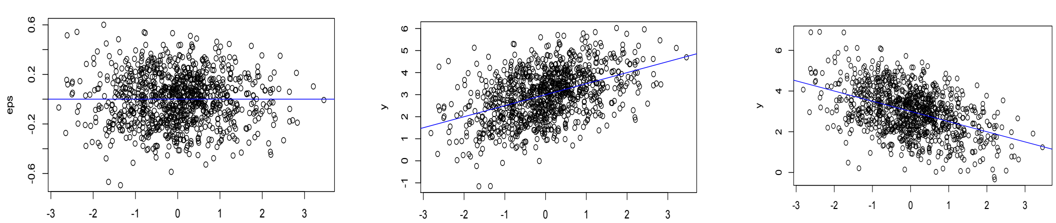
Correlation and covariance.

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.

$\rho = 0$ $\qquad\qquad$ $\rho = 0.44$ $\qquad\qquad$ $\rho = -0.44$.



Correlation and covariance.

Note also that cov(X,X) = var(X). Why?

cov(X,Y) = E(XY) – E(X)E(Y)

So cov(X,X) = E(X X) - μ^2.
And Var(X) = E(X^2) - μ^2.
So Cov(X,X) = Var(X).

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(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) = $E(X^2) - \mu^2 = 1/3 - ¼ = 1/12$.

cov(3X+Y, 4X-Y) = 12 (1/12) – 4
= -3.

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.

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

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

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

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

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

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

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.

We have basically done all of chapters 1-7.1. Ignore 6.7 and most of 6.3 on optimal play.