Last week's problems

For feedback, etc., commit your changes to GitHub and then include are reference to relevant issue (usually issue#1) in the committee message:

Or (for today) send R-scripts to me at: pmarjora@usc.edu

Reminder: This year, we are hosting the course on Github at https://github.com/PM520-Spring-2020

How the course will be examined

- Every few weeks you will be introduced to an examinable project or two.
- You have two weeks to work on it.
- You then have to turn in a written report and a script file containing your R code.
- I may run the code on test datasets.
- There will be 4-5 of these examinable projects.
- **There will also be a final in the form of small-group presentations:**
  - Short presentation about a topic we haven't covered in the course: e.g. E-M algorithms, fancy MCMC version, (e.g. Hamiltonian MCMC), Empirical Bayes methods, Boot-strapping,…), *or a related project from your own research.*
- Participation!
- **I don't this grade this class on a 'curve'.**

Grading: 70% projects + 20% final presentation + 10% participation

Github organization

- The repos for Week n will all have names that start with “Weekn-“.
- I have created an issue for each repository so that you can refer to that issue in your commit messages to tell me that there is something to look at or share with the class.
- I have created a separate “conversation” issue for each repository so that you can post messages to each other, or me, asking for thoughts, feedback, etc.
- I have created a repository called “Useful-Tips” for sharing useful tricks that you learn.
- All of the above will count towards “participation” in your final grade.
- The best form of participation is a willingness to share your code while it is ‘in development’ for comment/feedback, etc.

Examinable project1: Randomization tests - golf balls

- Allan Rossman used to live along a golf course and collected the golf balls that landed in his yard. Most of these golf balls had a number on them.

- Question: What is the distribution of these numbers?
- In particular, are the numbers 1, 2, 3, and 4 equally likely?

[Originally due to Allan Rossman - via Randall Pruim]

Examinable project 1: golf balls

- Population: Golf balls at that driving range
- Allan tallied the numbers on the first 500 golf balls that landed in his yard one summer.
- Sample: Those golf balls driven ~150 yards and sliced.

1	2	3	4
137	138	107	104

486 balls in total

- There were 14 “others”, which we will ignore
- Question: What is the distribution of these numbers? In particular, are the numbers 1, 2, 3, and 4 equally likely?
- Meta-Question: How do we answer this question using the data?

Randomization test set-up

- Null hypothesis: Our default belief about the data.
  – Here, it is that the numbers on the population golf balls are uniformly distributed between 1 and 4.
- Test statistic: A single number that can be calculated from the data and used to test whether our null hypothesis is true.

1	2	3	4
137	138	107	104

486 balls in total

- What test-statistic should we use?
- How should we conduct the test **using simulations**?

Examinable Project - what to turn in

- A 3-5 page write-up, in the form of a knitted file (ideally a .Md file) including the following:

- Explain the scenario you are asked to consider (golf balls)
- Explain the logic of a hypothesis test
  - 1. State Hypotheses
    - Null hypothesis that you must provide a model for (for simulations)
  - 2. Calculate a Test Statistic
  - 3. Determine the "p-value" - how do you do this via simulation?
  - 4. Interpret the p-value you get
- What makes a good test statistic? **Compare several**.
- Comment on things such as the following in a Conclusion section:
  - What would you do if you knew the sampling distribution of the test statistic? How does simulation help if you don't?
  - Power against particular types of alternatives:
  - Does the best test statistic depend on what you are testing?

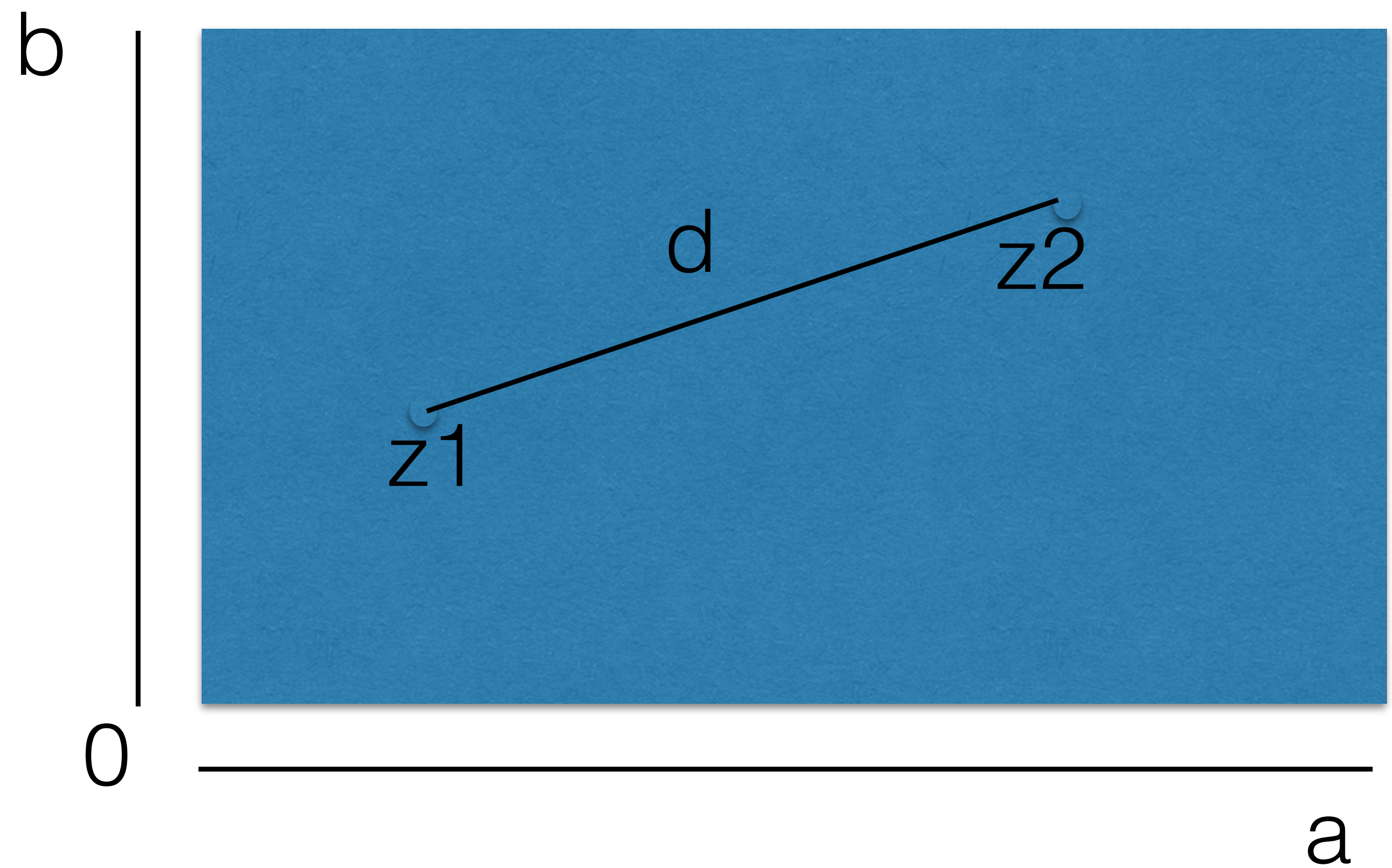
Examinable Projects - what to turn in

- What to turn in (due in one week at 1pm on Mon):
  - Write up a description of your method and then show the results.
  - Include a script showing your (clearly commented) code.
  - The whole thing should be about 3-5 pages
  - 1/2 page introduction - describe the idea of Monte Carlo simulation
    - 1-2 page of methods - describe the specific method you are using
    - 1-2 pages of results - show your results
    - 1/2 page of conclusions - summarize the results; what did you learn

– Submit both an Rmarkdown file and a 'knitted' file (i.e. a file that also shows the output). Alternatively, you can submit a knitted .Md or pdf, but pdfs can take a while to render on github.

Other Problems from last week: Distance between points

- Suppose we have a rectangle [0,a]x[0,b]
- If we generate two points, z1 and z2, randomly in the rectangle, what is the expected distance, d, between them?



Questions

- How many simulations should you run? (The program will return an answer for any value of NumberOfSims that you give it.)
-

$$\text{E(d)} = \frac{1}{15}\left[\frac{a^3}{b^2} + \frac{b^3}{a^2} + \sqrt{a^2+b^2}\left(3 - \frac{a^2}{b^2} - \frac{b^2}{a^2}\right)\right] + \frac{1}{6}\left[\frac{b^2}{a}\operatorname{arccosh}\left(\frac{\sqrt{a^2+b^2}}{b}\right) + \frac{a^2}{b}\operatorname{arccosh}\left(\frac{\sqrt{a^2+b^2}}{a}\right)\right]$$

where

$$\operatorname{arccosh}(t) = \log\left(t + \sqrt{t^2-1}\,\right)$$

Further questions

- For fixed area A, is the expected distance between two randomly chosen points, E(d), bigger for a square or a rectangle?

- Show that, for any fixed area A, and distance D, there exists a rectangle such that E(d)>D.

Monte Carlo application 2: Hypercubes

- Consider n-dimensional unit 'cubes'
  - n=3 -> a cube. Coordinates= (x_1,x_2,x_3) [or (x,y,z)].
  - n=2 -> a square. Coordinates= (x_1,x_2) [or (x,y)].
  - n=1 -> line. Coordinates= (x_1) [or (x)].

  - n=4 -> 4-dimensional hypercube. Coordinates= (x_1,x_2,x_3,x_4).
  - n=5 -> 5-dimensional hypercube. Coordinates= (x_1,x_2,x_3,x_4,x_5).
  - n=N -> N-dimensional hypercube. Coordinates= $(x_1,x_2,x_3,x_4,\ldots,x_N)$.

- Let's suppose that all are unit cubes, so $0 \leq x_i \leq 1$, for all i.
- **Question: what proportion of the volume of an N-dimensional hypercube is within a distance of 0.1 of the surface?**

Question: If we sample a point ***uniformly at random*** from inside a hypercube, what is the probability that the point is within a distance of 0.1 of the surface?

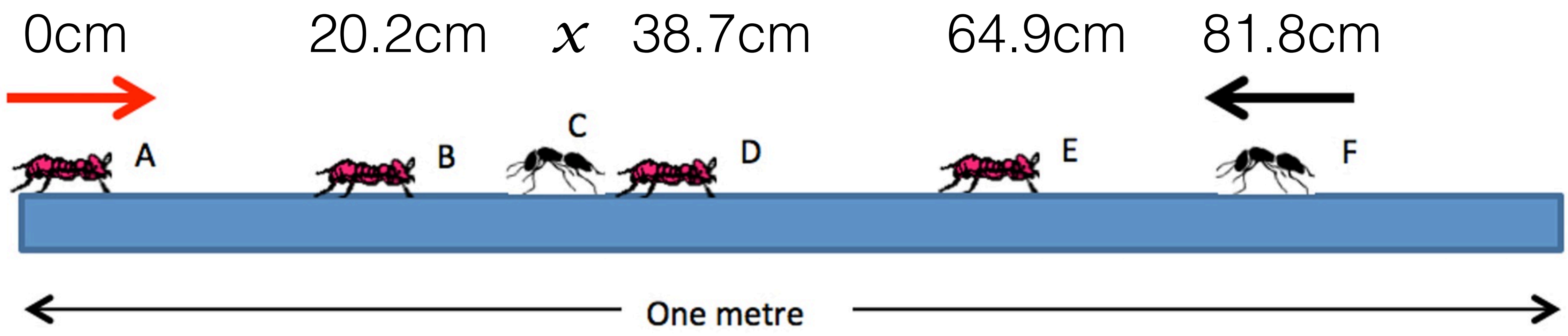
Monte Carlo Application 3: coin tossing

- Suppose we simulate a sequence of 500 coin tosses. Define a ‘run’ as a sequence of all Heads (and of all Tails). Answer the following:
  - What is the distribution of the number of heads?
  - What is the distribution of the length of the run starting at the first toss?
  - What is the expected number and distribution of the number of runs in total?
  - What is the expected value and distribution of the length of the longest run?

MC Application 3: Answers

- Suppose we simulate a sequence of n coin tosses. Define a 'run' as a sequence of all Heads (or all Tails). Answer the following:
  - What is the distribution of the number of heads?
    - Binomial(n,1/2)
  - What is the distribution of the length of the run starting at the first toss?
    - Geometric(0.5)
  - What is the expected number and distribution of the number of runs in total?
    - 1+Binomial(n-1,1/2)
  - What is the expected value and distribution of the length of the longest run?
    - ?

Ants on a stick



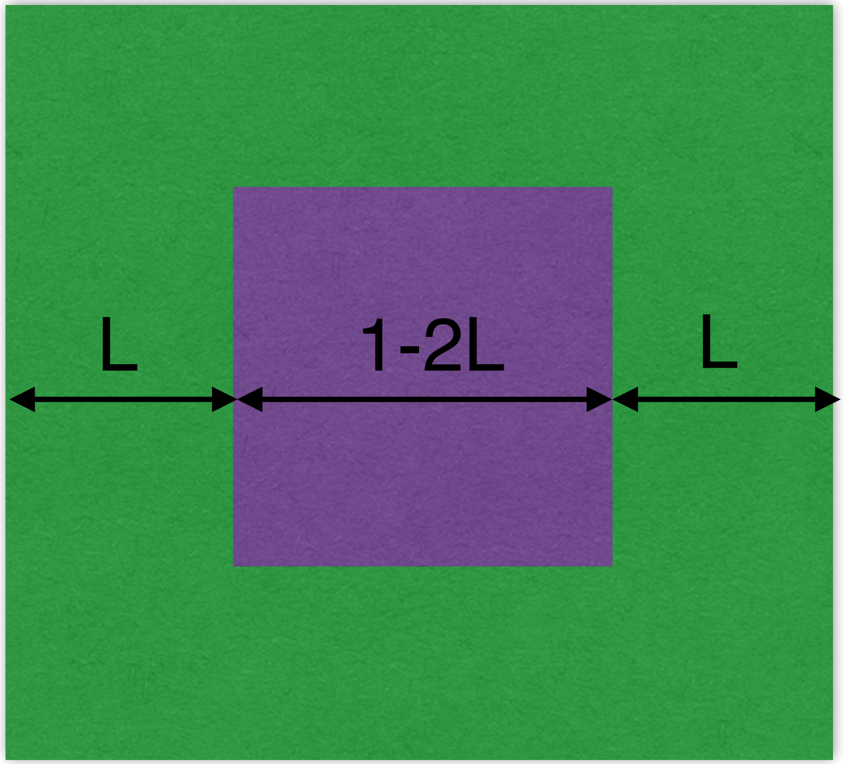
- Ants walk at 1cm/second. When they meet, each ant turns around and walks in the other direction. When they reach the end of the stick they fall off.
  - How many seconds until the last ant falls off?
  - Which ant is the last to fall off the stick?

In-class exercise:

- Estimate (not calculate) π using Monte Carlo methods.
- So, you need to think of something you can simulate in which the probability of success depends upon π.

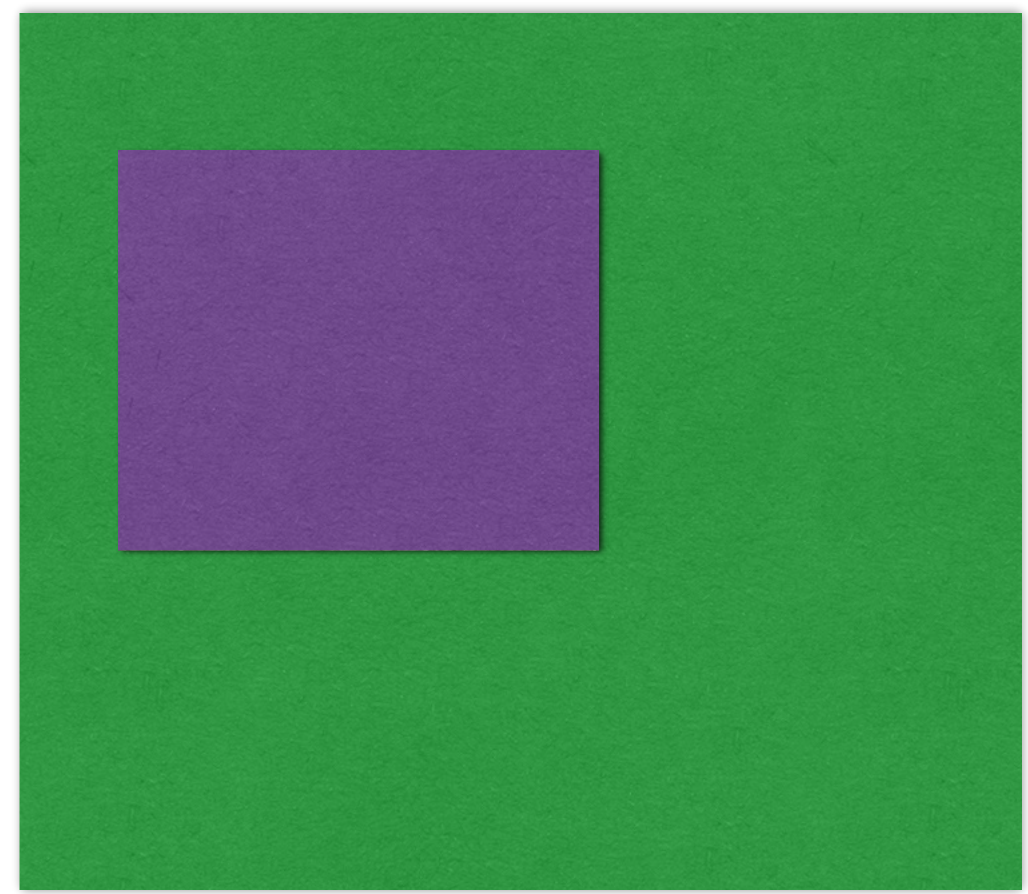
Hypercubes

- Answer:
  - The proportion of the volume that is within a distance of L of the surface of an N-dimensional hypercube with sides of length 1 is …. $1-(1-2L)^N \rightarrow 1$ as $N \rightarrow \infty$ (for $0 \leq L \leq 0.5$).

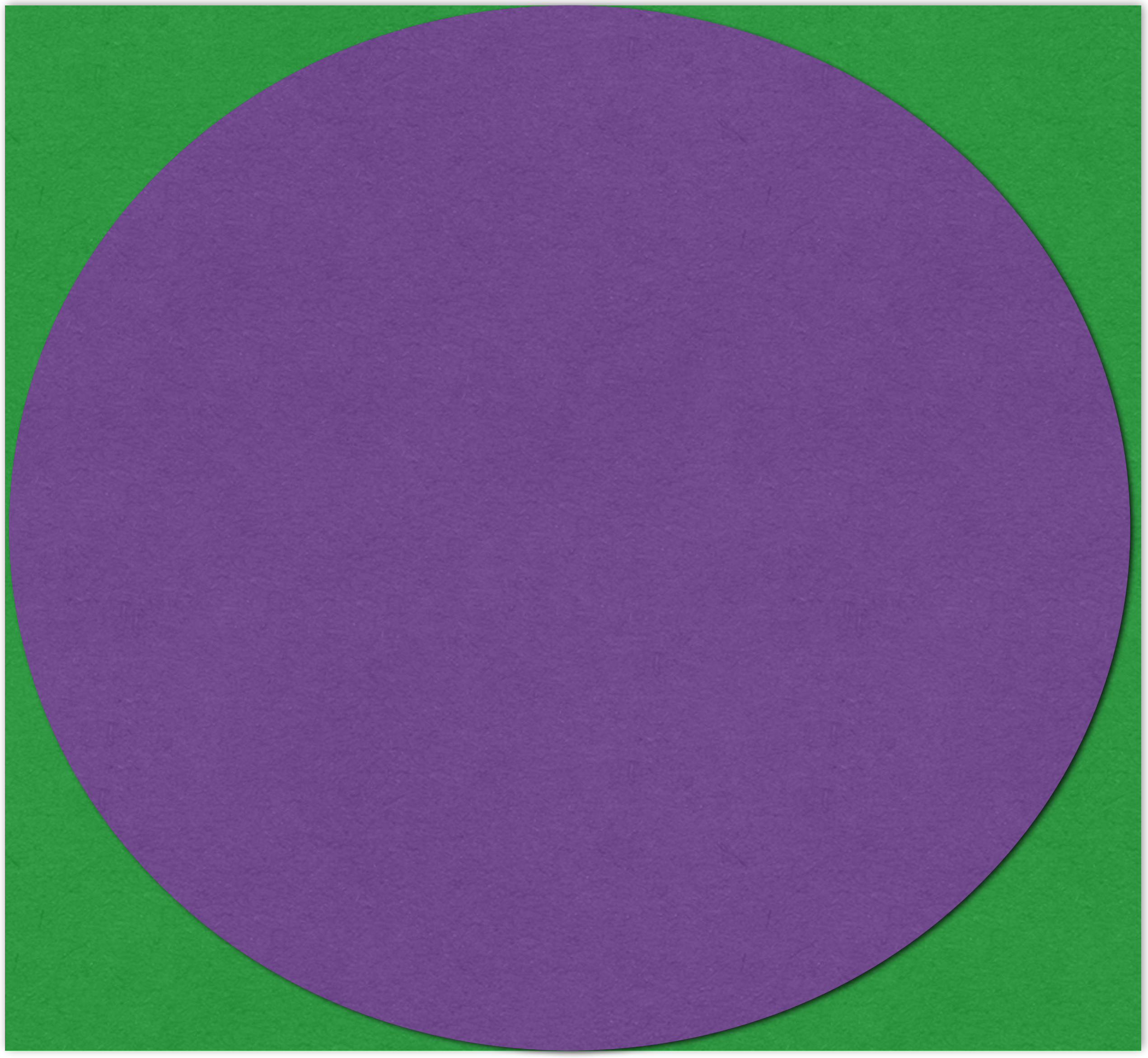


Hint

- Suppose we randomly pick a point in the green shape.
- What is the probability that the point lands inside the purple shape?



- Answer:
  – (purple area)/(green area)



Jocky Wilson - Athlete, legend.

"Jocky Wilson is the Citizen Kane of darts." - Guardian newspaper

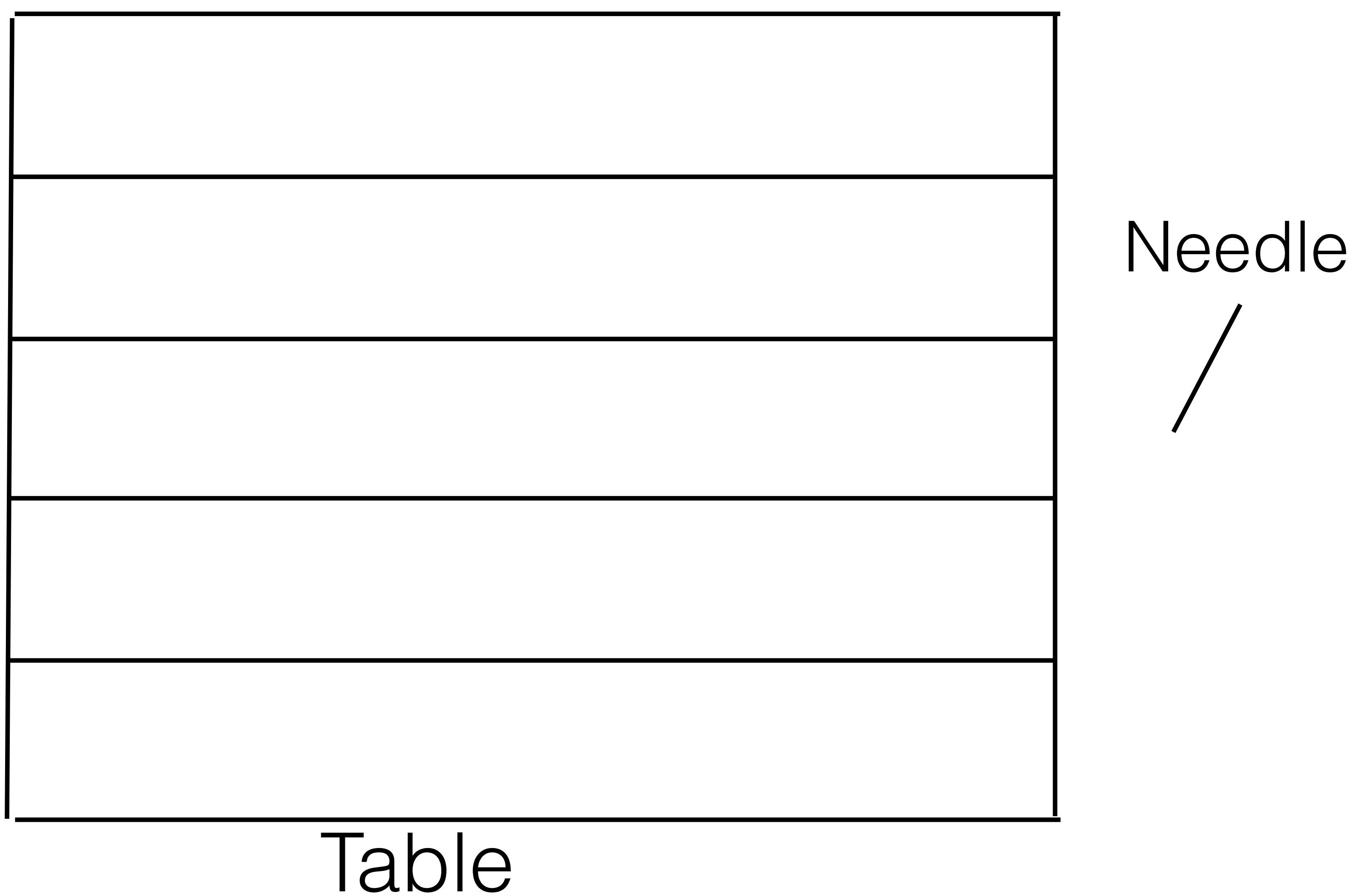
In 1989 he released a record "Jocky on the oche" but it failed to spark the public imagination and is reputed to have sold just 850 copies. (Wikipedia).

- As he mourned him yesterday, old rival and pal Bobby George called him “the king of darts”.

“In any case, a lot of Jocky’s diet was liquid. He liked to prepare for a game with “seven or eight vodkas, to keep my nerves”. Jocky also knocked back pints of lager during matches, which didn’t always help his game.

http://www.dailyrecord.co.uk/news/editors-choice/2012/03/26/pals-and-stars-honour-scots-darts-legend-jocky-wilson-after-he-dies-at-home-aged-62-86908-23801861/

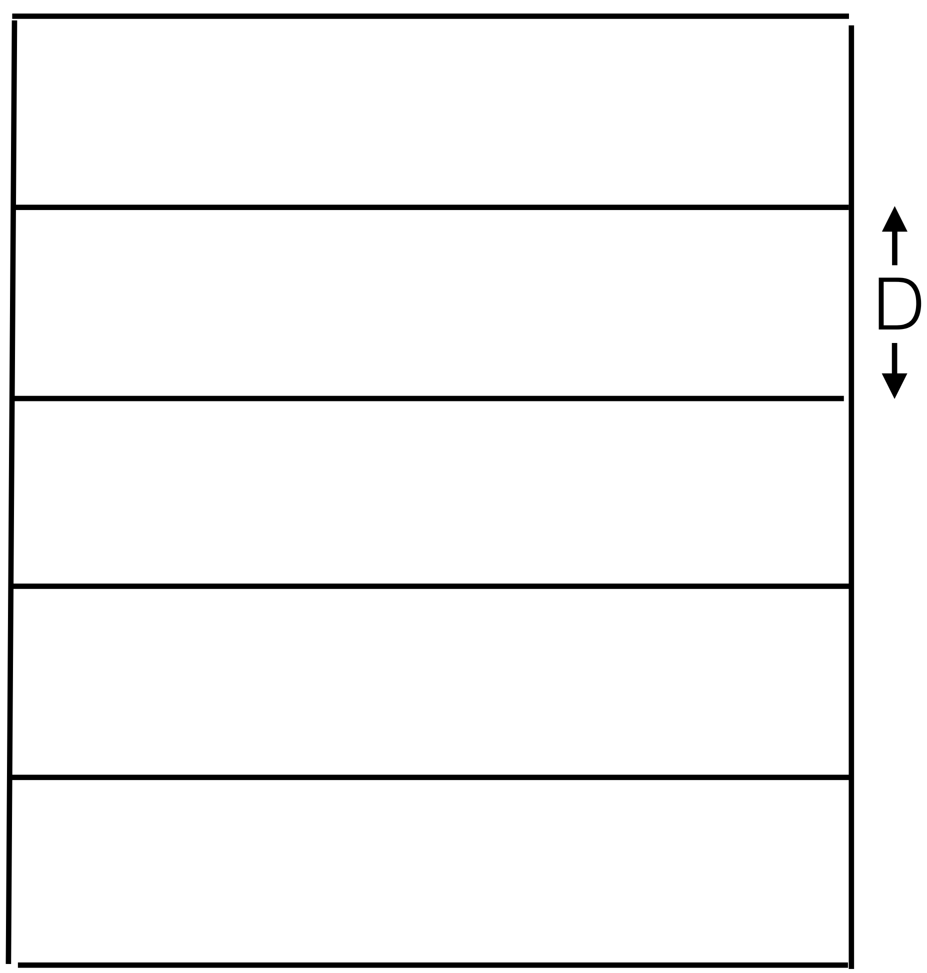
Monte Carlo Simulation: Calculating π - Buffon’s needle



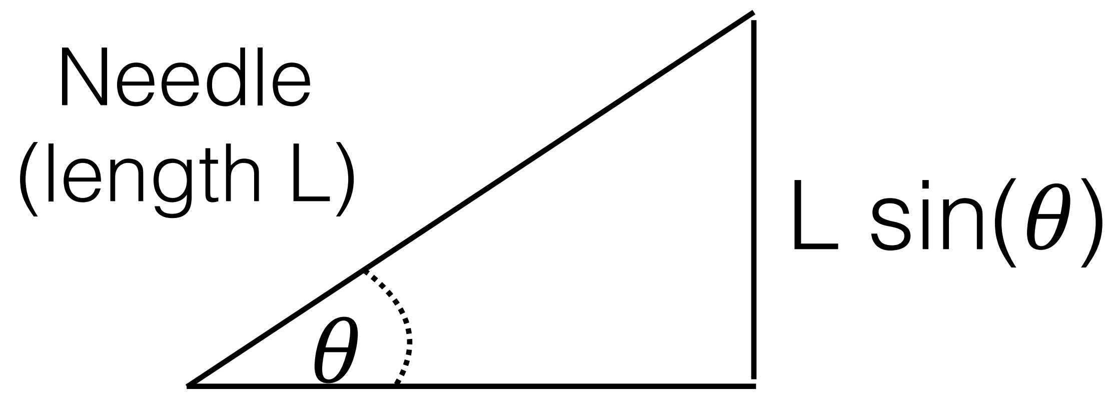
Georges-Louis Leclerc, Comte de Buffon
(French pronunciation: [ʒɔʁʒ lwi ləklɛʁ kɔ̃t də byfɔ̃]; 7 September 1707 – 16 April 1788)

- http://www.angelfire.com/wa/hurben/buff.html

Monte Carlo Simulation: Calculating π - Buffon's needle



Table



So, P(cross line) = $\min(1, L\sin(\theta)/D)$

If N= number of trials, lines are distance **D=1** apart,
C = number of times needle crosses line,
and **L≤D**, then:
$C/N \sim \int L\sin(\theta)(1/\pi)\, d\pi$ [integral goes from 0 to π in radians]
So $C/N \sim L[-\cos(\pi)-(-\cos(0))]/\pi = 2L/\pi$
So $\pi \sim 2LN/C$

Results:

- Example:
  - Mario Lazzerini (1901):
    - Needle length: 2.5cm
    - Distance between lines: 3cm
    - 3408 throws
    - 1809 hits of the lines
    - Led to estimate of pi of 3.1415929
  - In order to guarantee accuracy to 6 decimal places would need 134 trillion needles.
- What is the optimal length of needle?

Classical statistical inference

- H_o: $\theta=1$ versus H_a: $\theta>1$
- Classical approach
  - Calculate a test statistic
  - P-value = P(Test stat. value (or “more extreme”) | $\theta=1$)
  - P-value is NOT P(Null hypothesis is true)
  - Confidence interval [a, b] : What does it mean?
- But scientist wants to know:
  - P($\theta=1$ | Data)
  - P(H_o is true) = ?
- Problem
  - θ “not random”

Bayes Theorem

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$= \frac{P(B \mid A)P(A)}{P(B)}$$

$$= \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A^C)P(A^C)}$$

Thomas Bayes was an English statistician, philosopher and Presbyterian minister who is known for having formulated a specific case of the theorem that bears his name: Bayes' theorem. Wikipedia
Born: 1702, London, United Kingdom
Died: April 7, 1761, Royal Tunbridge Wells, United Kingdom
(wikipedia)

Example Application of Bayes' theorem

- Population has 10% liars
- Lie Detector gets it “right” 90% of the time.
- Let A = {**Actual** Liar},
- Let L = {Lie Detector reports you are **Liar**}
- Lie detector reports suspect is a liar. What is probability that suspect actually is a liar?

$$P(A \mid L) = \frac{P(L \mid A)P(A)}{P(L \mid A)P(A) + P(L \mid A^C)P(A^C)}$$

$$= \frac{(.90)(.10)}{(.90)(.10) + (.10)(.90)} = \frac{1}{2}!!!!!$$

Bayesian statistics

- Paradigm shift in statistical philosophy
  - θ assumed to be a realization of a random variable
  - Allows us to assign a probability distribution for θ based on *prior* information
  - 95% "confidence" interval $[1.34 < \theta < 2.97]$ means what we "want" it to mean: e.g., $P(1.34 < \theta < 2.97) = 95\%$

Bayesian modeling/statistics

- Three General Steps for Bayesian Modeling
  - I. Specify a probability model for unknown parameter values, that includes some prior knowledge about the parameters (if available).
  - II. Update knowledge about the unknown parameters by conditioning this probability model on observed data using Bayes' Theorem.
  - III. Evaluate the fit of the model to the data *and the sensitivity of the conclusions to the assumptions (i.e. the prior)*.

Bayesian statistics

- Let θ represent parameter(s)
- Let X represent data

$$f(\theta \mid X) = f(X \mid \theta) f(\theta) / f(X)$$

- Left-hand side is a function of θ
- Denominator on right-hand side does not depend on θ

$$f(\theta \mid X) \propto f(X \mid \theta) f(\theta)$$

- Posterior distribution $\propto$ Likelihood x Prior distribution
- Posterior dist'n = Constant x Likelihood x Prior dist'n
- Goal: Explore the posterior distribution of θ

Prior distributions

- The prior distribution reflects what you knew about θ, the model parameter(s), before you did the experiment.

- Where do priors come from?
  - Previous studies, published work.
  - Researcher intuition.
  - Substantive Experts
  - Convenience (conjugacy, vagueness).

Simple example

- 'Biased coin' estimation: *P(*Heads*)* = *p* = *?*
- $X_1, \ldots, X_n$ 0-1 i.i.d. Bernoulli(*p*) trials
- Let X be the sequence of 'heads' and 'tails' in *n* trials
- Likelihood is $f(X \mid p) = p^X (1-p)^{n-X}$
- For prior distribution, could use *uninformative* prior
  - Uniform distribution on (0,1): *f(p)* = 1
- So posterior distribution is proportional to

  $f(X|p)f(p) = p^X (1-p)^{n-X}$

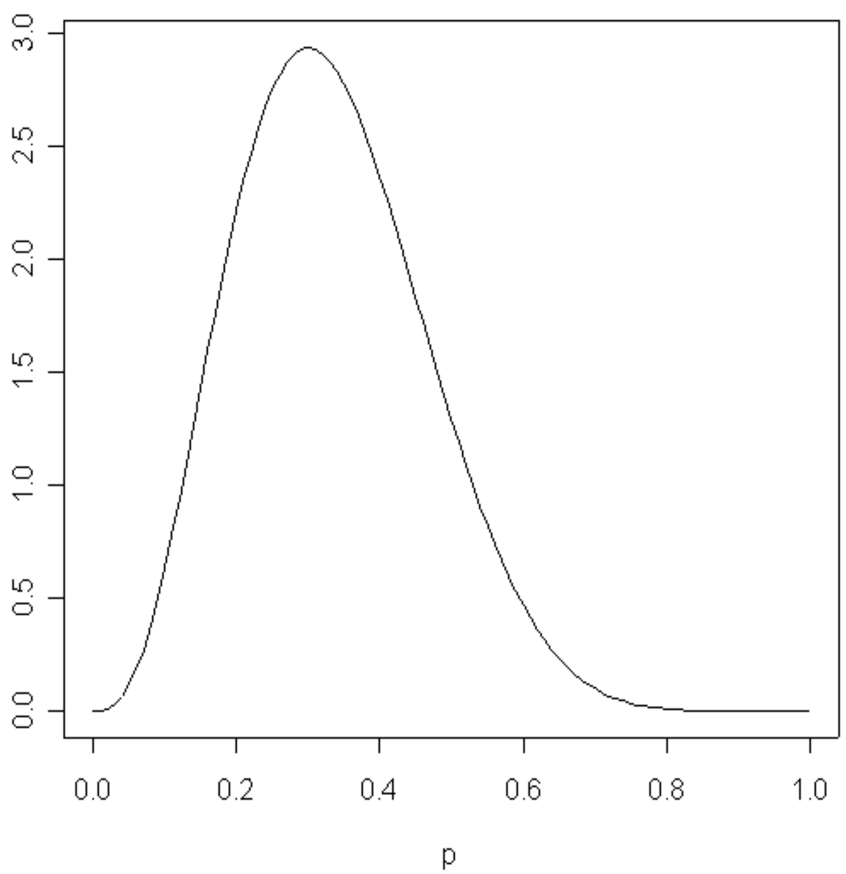
- $f(p|X) \propto p^X (1-p)^{n-X}$

Simple example (continued)

- Posterior density of the form $f(p)=Cp^x(1-p)^{n-x}$
- Beta distribution: Parameters x+1 and n-x+1
- Note that the Beta(1,1) is a Uniform(0,1) distribution.

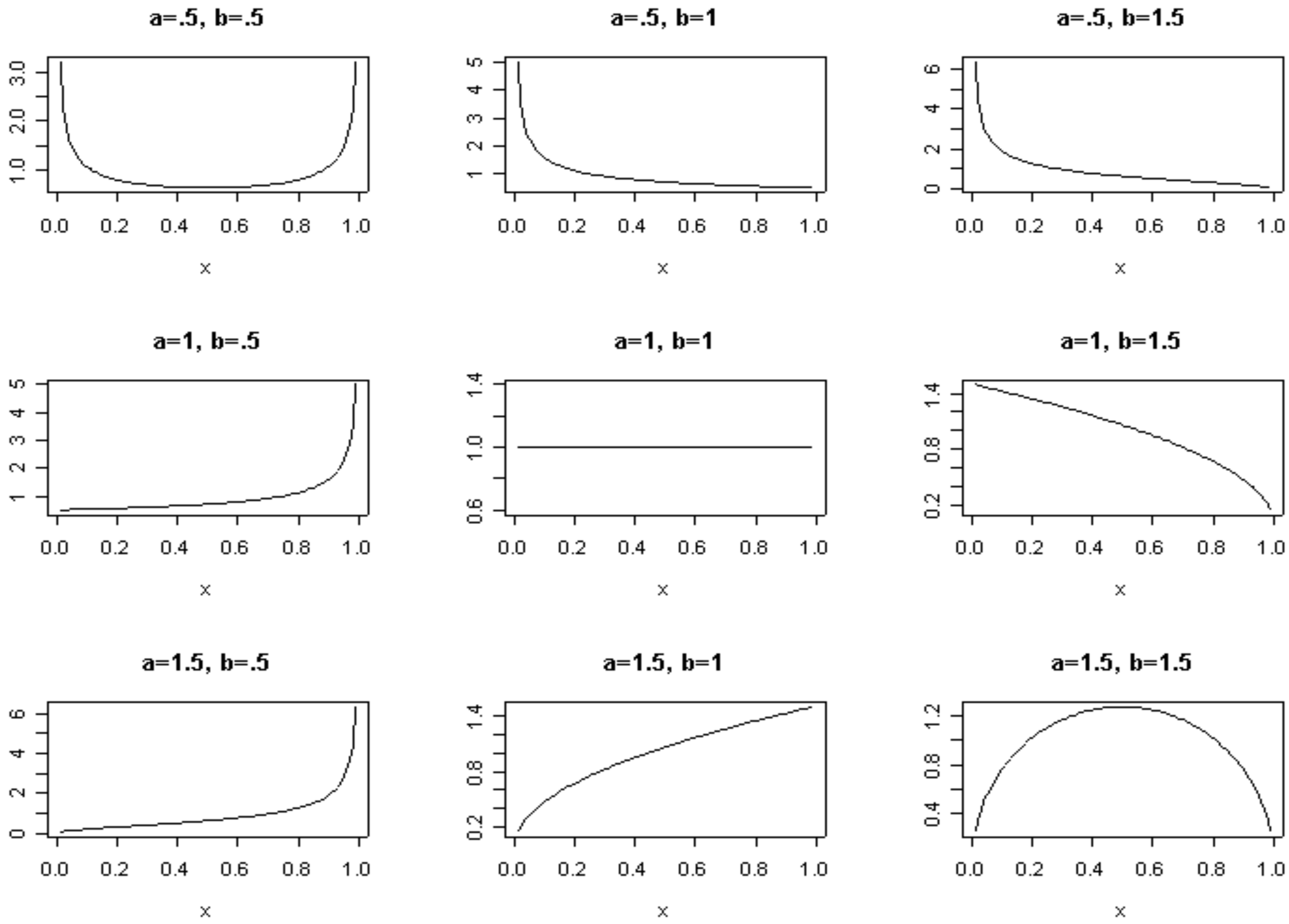
- *Example:* Data: 0, 0, 1, 0, 0, 0, 0, 1, 0, 1
- *n=10*
- *Use uniform [Beta(1,1,)] prior*

- Posterior dist'n is Beta(3+1,7+1) = Beta(4,8)
  - Mean: 0.33
  - Mode: 0.30
  - Median: 0.3238
  - 95% *credible* interval for *p is [0.11, 0.61]*
    - P(0.11 < *p* < 0.61 | *X*) = 0.95



Choice of prior

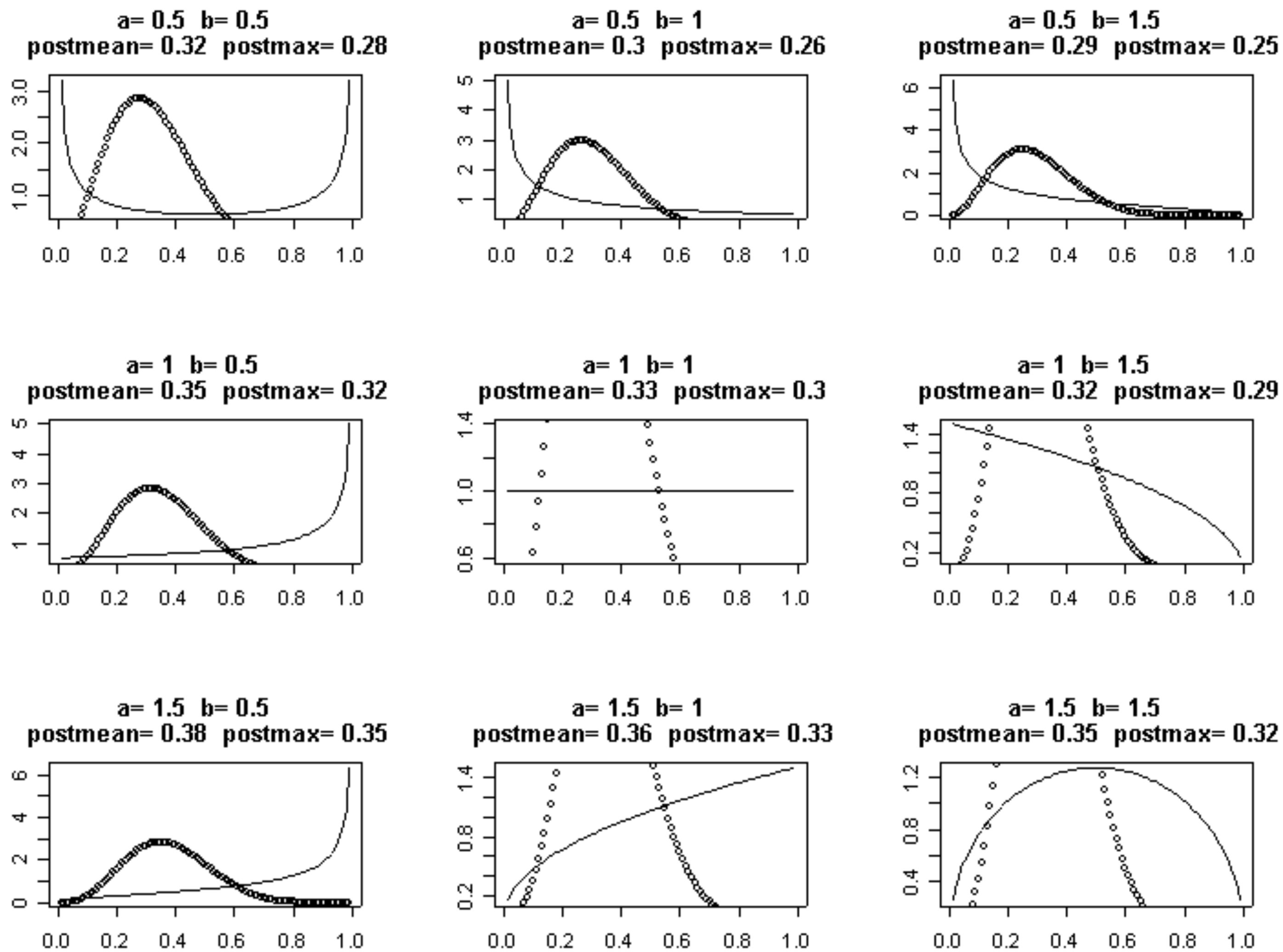
- Could use some other Beta distribution:



  - http://www.people.carleton.edu/~rdobrow/courses/275w05

Choice of prior

- Would get these posterior dist'ns:



[Priors that result in posteriors of the same form (i.e. same distributional family) are called *conjugate priors*.]

- http://www.people.carleton.edu/~rdobrow/courses/275w05

Differences Between Bayesians and Frequentists

- Frequentist:
  - The parameters of interest are fixed and unchanging under all realistic circumstances.
  - No information prior to the model specification.

- Bayesian:
  - View the world probabilistically, rather than as a set of fixed phenomena that are either known or unknown.
  - Prior information abounds and it is important and helpful to use it. *But results are now sensitive to priors.*

http://www.stat.ufl.edu/archived/casella/Talks/BayesRefresher.pdf

Generation of random variables, probability distributions, etc. (chapters 14/18)

- R, random number function:

```
runif(n,x,y)
```

- Generates n continuous random numbers distributed uniformly between x and y. e.g.,

```
> runif(10,0,5)
    [1] 0.7615195 2.6318839 1.4084045 1.3250196 3.6335061 4.8151777 2.7537802 2.7817826 1.3949205 0.7280294
```

To turn that into integers use the ceiling() function (ceiling(x) returns the smallest integer bigger than x):

```
> WhichBall<-ceiling(runif(10,0,NumberOfBalls))
> WhichBall
 [1]  7  7  3 10  3  3  2  2 10  2

> sample(x, size, replace = FALSE, prob = NULL)
> sample(1:10,10,TRUE)
 [1]  9  4  7  2  2  6 10  8  6  1
```

Sampling from Distributions - Discrete random variables

- Random variable [rv], X, takes a value x ∊ Ω (the state-space).
- e.g. X ~ coin toss: Ω = {"Head", "Tail"}
- e.g. Y ~ roll 6-sided die: Ω = {1,2,3,4,5,6}.
- In R, e.g., sample(c("H","T"),5,replace=TRUE)

- P(X=x) is a map from Ω to the unit interval [0,1], that gives the probability of the outcome x. [N.B., rv= capital letter; outcome= lower-case letter.]
- e.g. P(X="head") = 1/2
- e.g. P(Y=5) = 1/6.

Sampling from Distributions - Discrete random variables

- $F(x) = P(X \leq x)$ is the *Cumulative Distribution Function*.
- $F(x) = \sum_{y \leq x} P(X=y)$, for discrete random variables
- It follows that $P(a<X \leq b)=F(b)-F(a)$.

Sampling from Distributions - Continuous random variables

- $F(y)=P(Y\leq y)=\int_{-\infty}^{y} f(u)du$ [the Cumulative Distribution Function [or CDF]].
- $f(y)=dF(y)/dy$, is the *probability density function.*
  - e.g. exponential distribution

    $f(x) = \lambda \exp(-\lambda x), \quad x \geq 0$

    $F(x) = P(X<x) = 1-\exp(-\lambda x), \quad x \geq 0 \quad [0 \leq F(x) \leq 1]$

Empirical Density Estimates of Probabilities for Discrete rvs

- Simulate N, independent and identically distribution random variables, $X_1,X_2,\ldots,X_N \sim X$
- $f(x) = P(X=x) \simeq \sum_{i=1,\ldots,N} \mathrm{I}(X_i=x)/N$
- $F(x) = P(X\leq x) \simeq \sum_{i=1,\ldots,N} \mathrm{I}(X_i\leq x)/N$

where I is an indicator variable (takes the value 1 if true; 0 otherwise)

Empirical Density Estimates of Continuous Variables

- Simulate N, independent and identically distributed random variables, $X_1, X_2, \ldots, X_N \sim X$
- $F(x) \simeq \Sigma_{i=1,\ldots,N} \mathrm{I}(X_i \leq x)/N$

  where I is an indicator variable (1 if true; 0 otherwise)

- Cannot estimate f(x) using the same strategy as for discrete random variables (why?). Instead, we will (informally speaking) use histograms to estimate f(x).

Simulating discrete random variables - Chapter 18

- To generate a random variable X=x, with density f() and cdf F():
- Sample u from F(x) [i.e., sample u from Unif[0,1]].
- x= $F^{-1}(u)$. (discrete r.v.: find the smallest x such that u ≤ F(x))
- e.g.

```
set.seed(1473)
# sample from Unif[0,1,]
u<-runif(1,0,1)

# sample X from some distribution F (on the non-negative integers, say)
X<-0
while (F(X)<u){
   X <- X+1
   }
```

Example: Binomial random variables (c.f. page 335 of text)

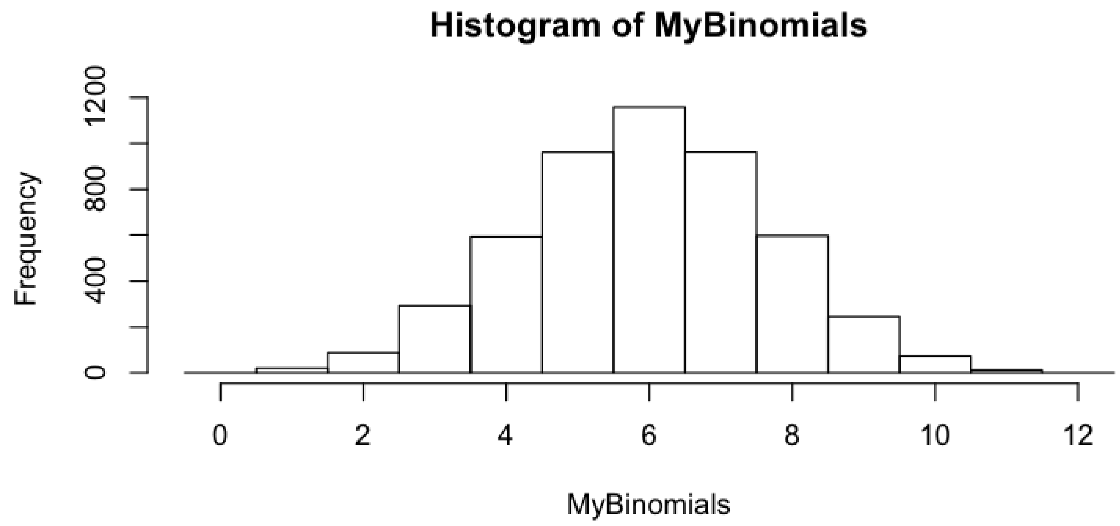
```
set.seed(1473)

binom.cdf<-function(x,n,p){
	Fx<-0
	for (i in 0:x){
		Fx <- Fx + choose(n,i)*p^i*(1-p)^(n-i)
	}
	return (Fx)
}
cdf.sim<-function(F,...){
	X <- 0
	U <- runif(1) # defaults to bounds of 0 and 1
	while (F(X,...)<U){
		X <- X+1
	}
	return (X)
}

MyBinomials<-numeric()
for (i in 1:5000){
		MyBinomials[i]<-cdf.sim(binom.cdf,12,0.5)
}

MyBreaks<-seq(0,13,1)
MyBreaks<-MyBreaks-0.5
BinHist<-hist(MyBinomials,breaks=MyBreaks)
```

... = unspecified number of other arguments



[In repo 'Binomials' onGithub]

Continuous Random Variable Example: Exponential random variables

- $f(x) = \lambda \exp(-\lambda x)$
- $F(x)=P(X<x)= 1 - \exp(-\lambda x)$ $[0 \le F(x) \le 1]$
- $u \sim U[0,1] \sim F(x)$. So $x \sim F^{-1}(u)$.

- Set $u=F(x)=1-\exp(-\lambda x)$

  So $\exp(-\lambda x)=1-u$

  and $x=(-1/\lambda)\log(1-u)=F^{-1}(u)$

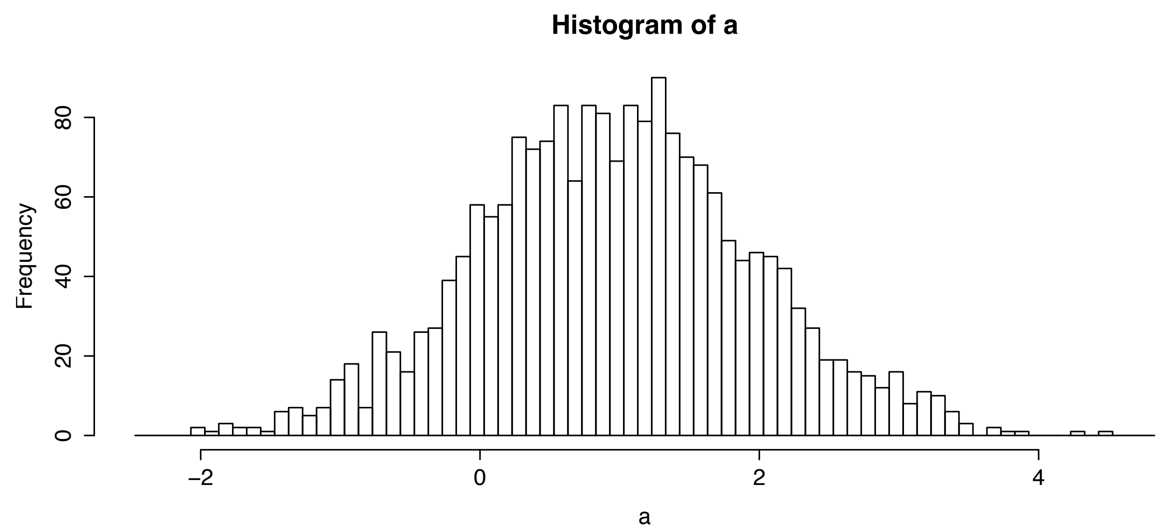
```
lambda <- 1.1   # the (example) parameter for the exponential distn
u <-runif(1000,0,1)
ExpRVs <- (-1/lambda)*log(1-u)  # note that this works even when u is a vector
hist(ExpRVs)
# or if U~U(0,1), so is 1-U, so....
lambda <- 1.1   # the parameter for the exponential distn
u <- runif(1000,0,1)
ExpRVs <- (-1/lambda)*log(u)
hist(ExpRVs)
```

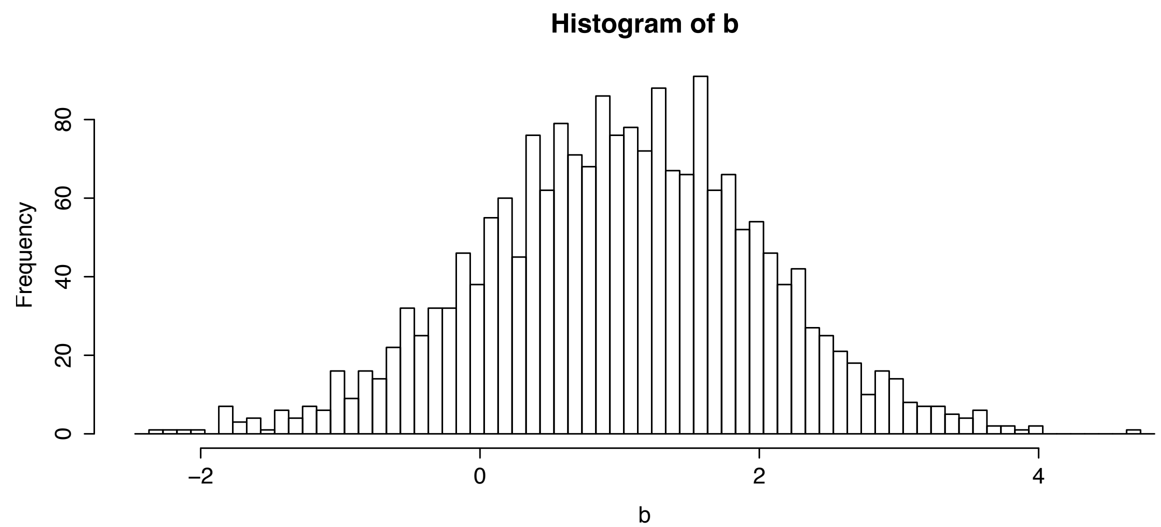
QQ plots

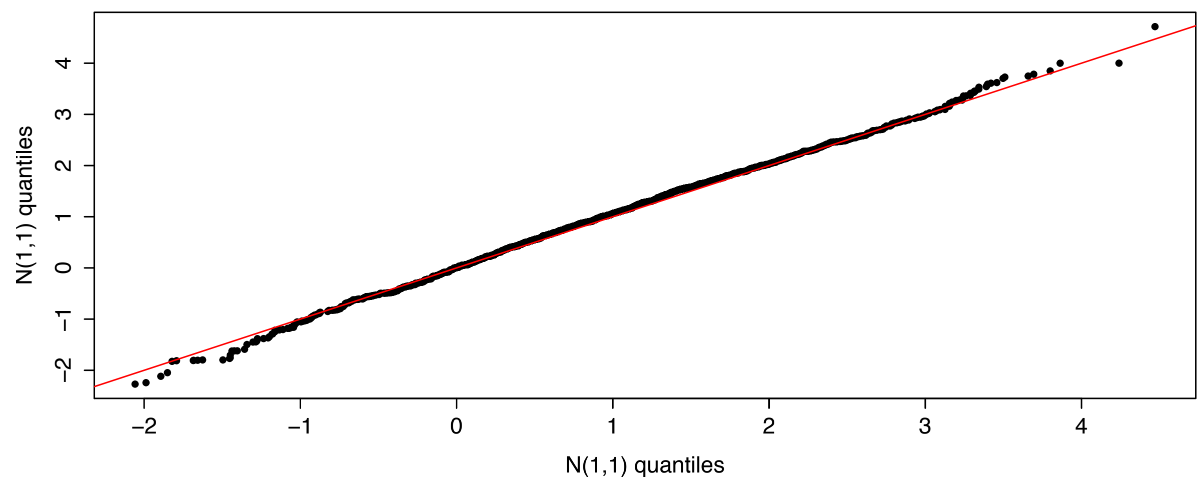
- Used to compare samples, $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$:
  - Order the data points in each sample from low to high, to get $X_{[1]},\ldots,X_{[n]}$ and $Y_{[1]},\ldots,Y_{[n]}$
  - Plot $X_{[1]}$ against $Y_{[1]}$, $X_{[2]}$ against $Y_{[2]}$, $X_{[3]}$ against $Y_{[3]}$, etc.
  - If the distributions are the same, you should see a straight line (for large samples)
  - 'qqplot' in R

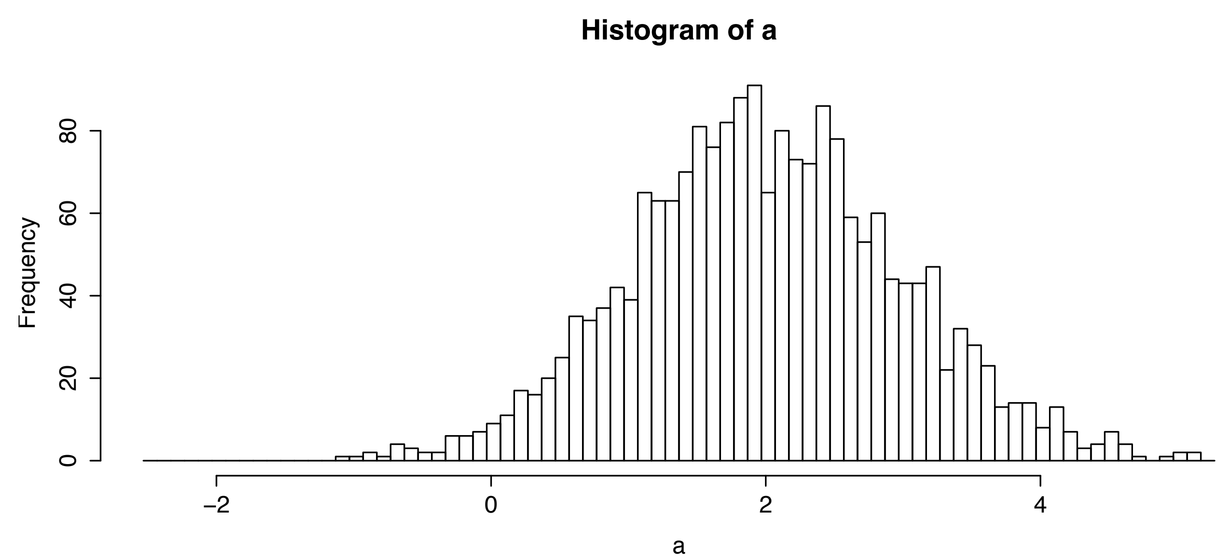
- Can do the same with one sample and Normal random deviates (qqnorm in R)

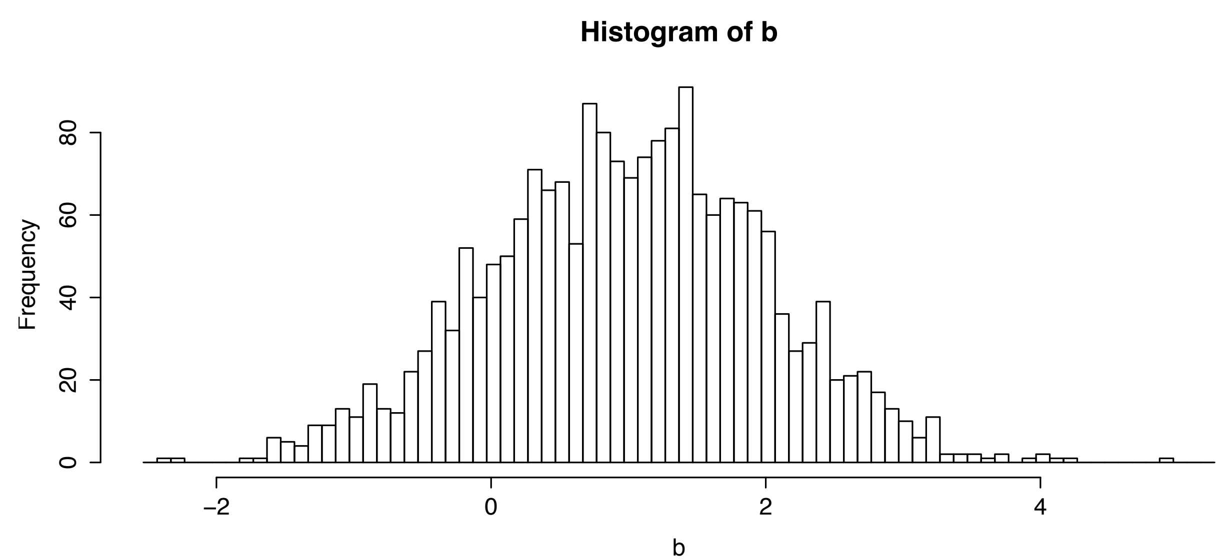
- Formal tests: Kruskal-Wallis test or ANOVA.

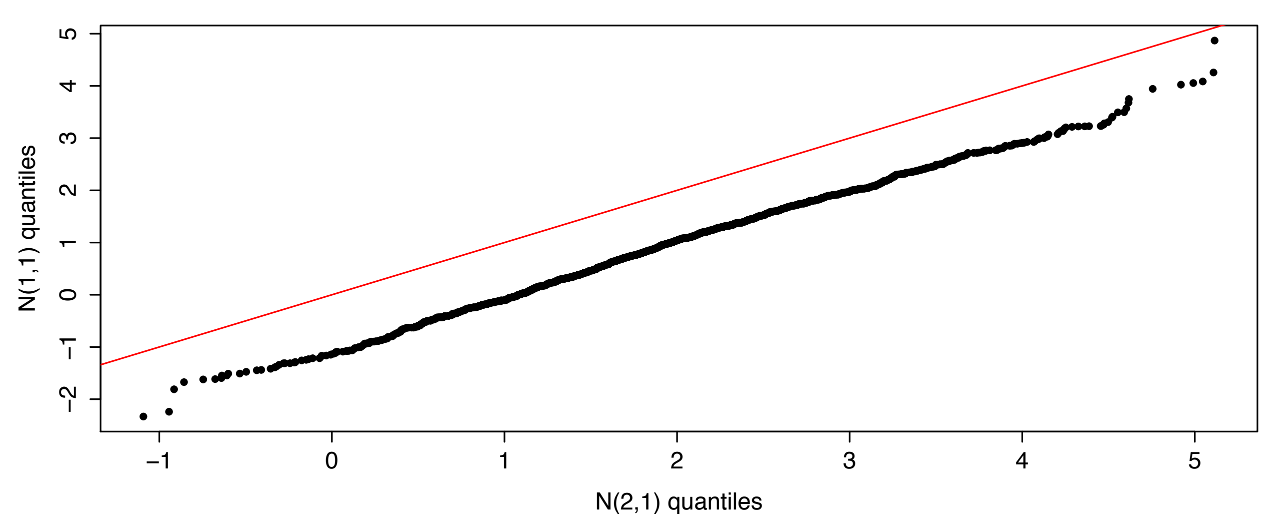

Histogram of a
Frequency
0
20
40
60
80
−2
0
2
4
a

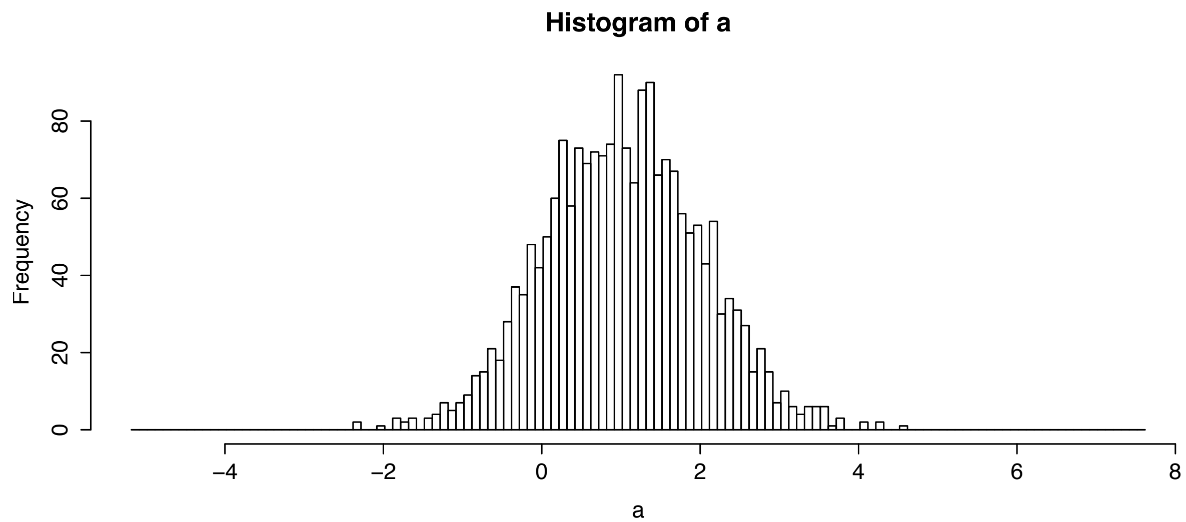

Histogram of b
Frequency
0
20
40
60
80
−2
0
2
4
b

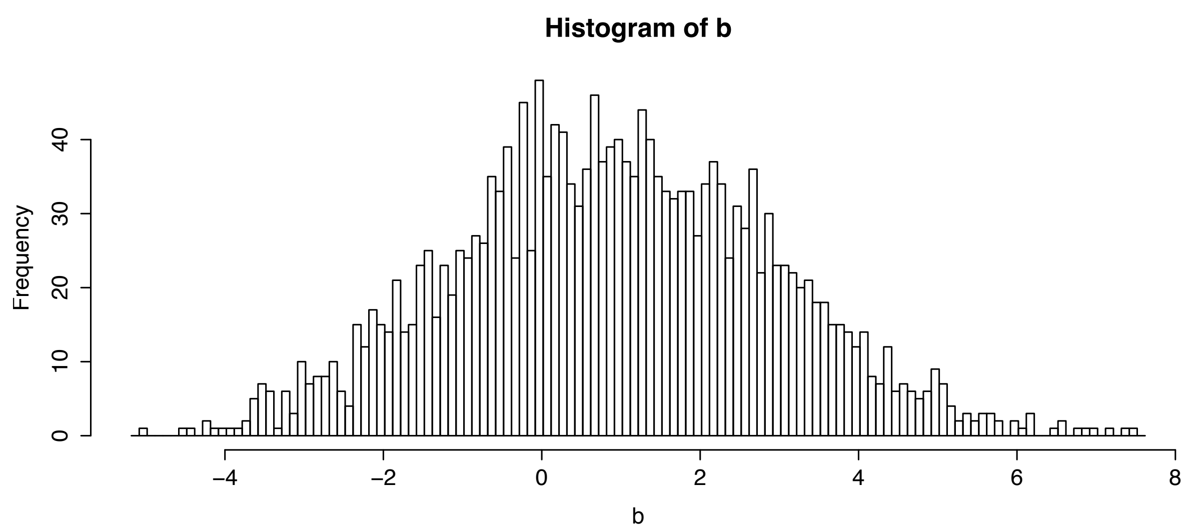

N(1,1) quantiles
−2
−1
0
1
2
3
4
N(1,1) quantiles

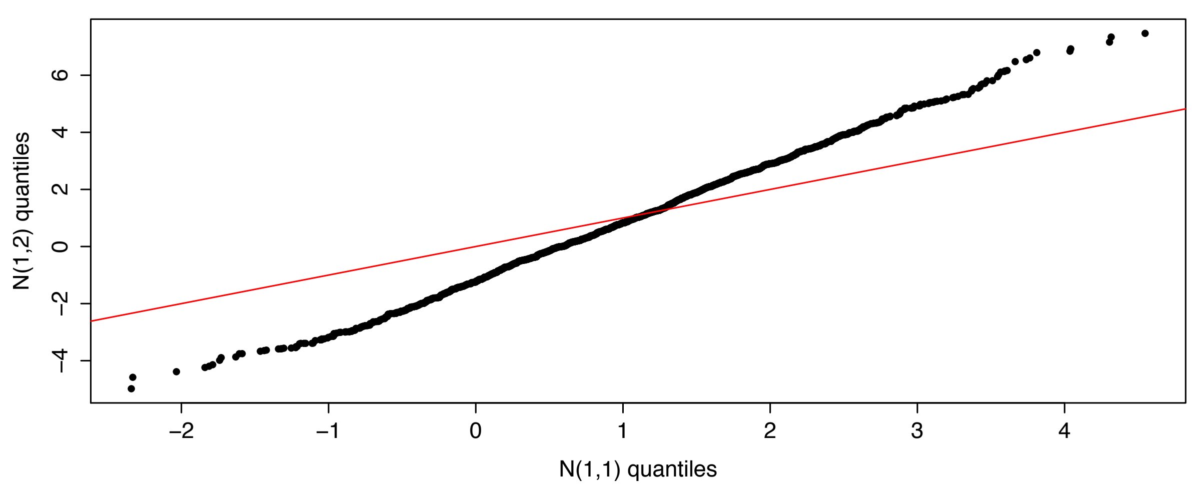

Histogram of a
Frequency
0
20
40
60
80
−2
0
2
4
a

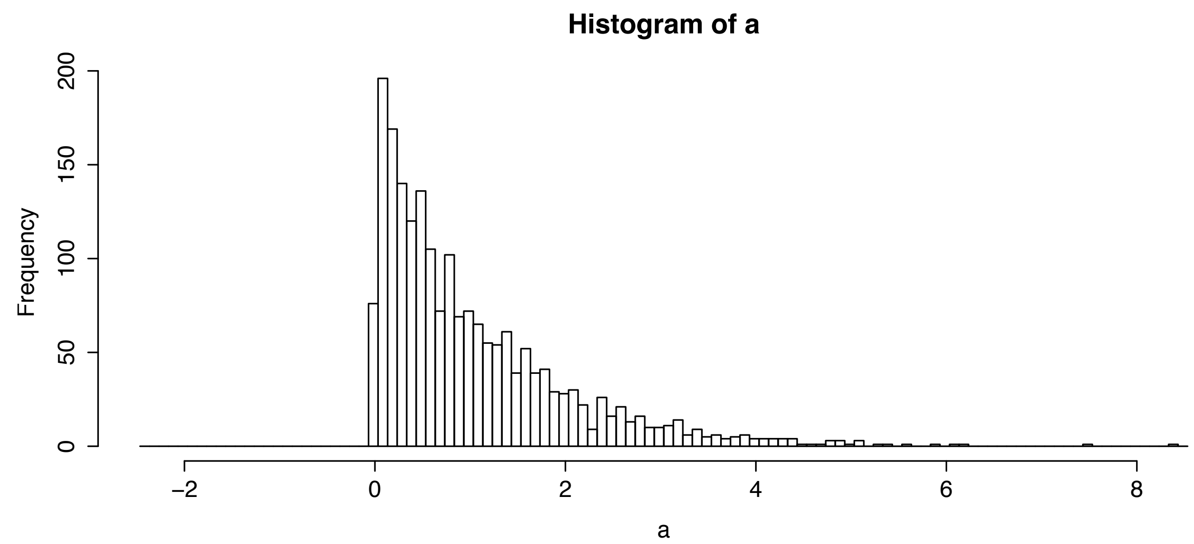

Histogram of b
Frequency
0
20
40
60
80
−2
0
2
4
b


N(1,1) quantiles
−2
−1
0
1
2
3
4
5
−1
0
1
2
3
4
5
N(2,1) quantiles

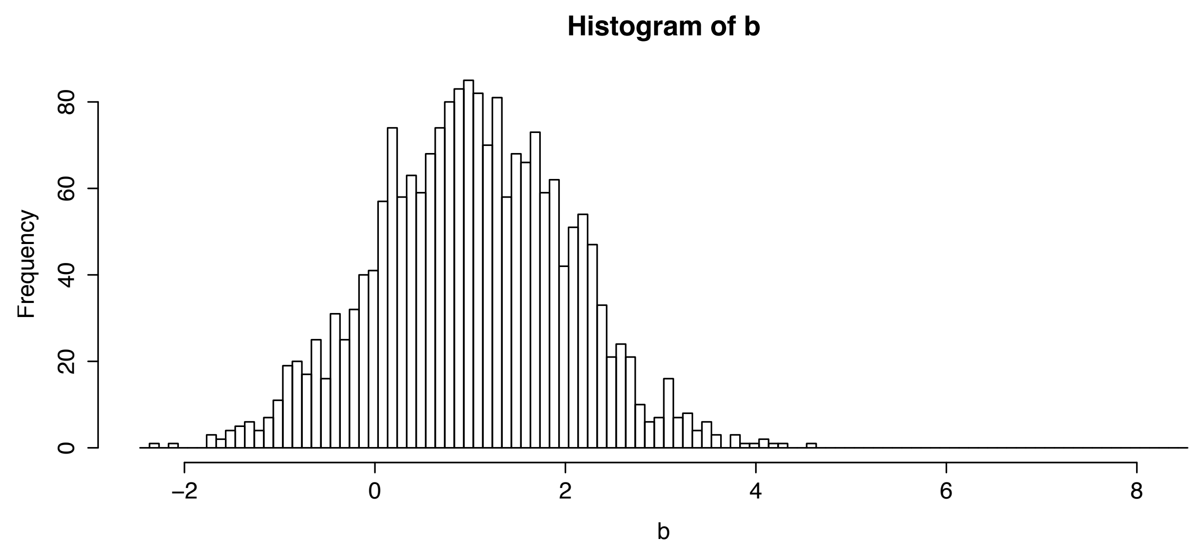

Histogram of a
Frequency
0
20
40
60
80
−4
−2
0
2
4
6
8
a


Histogram of b
Frequency
0
10
20
30
40
−4
−2
0
2
4
6
8
b

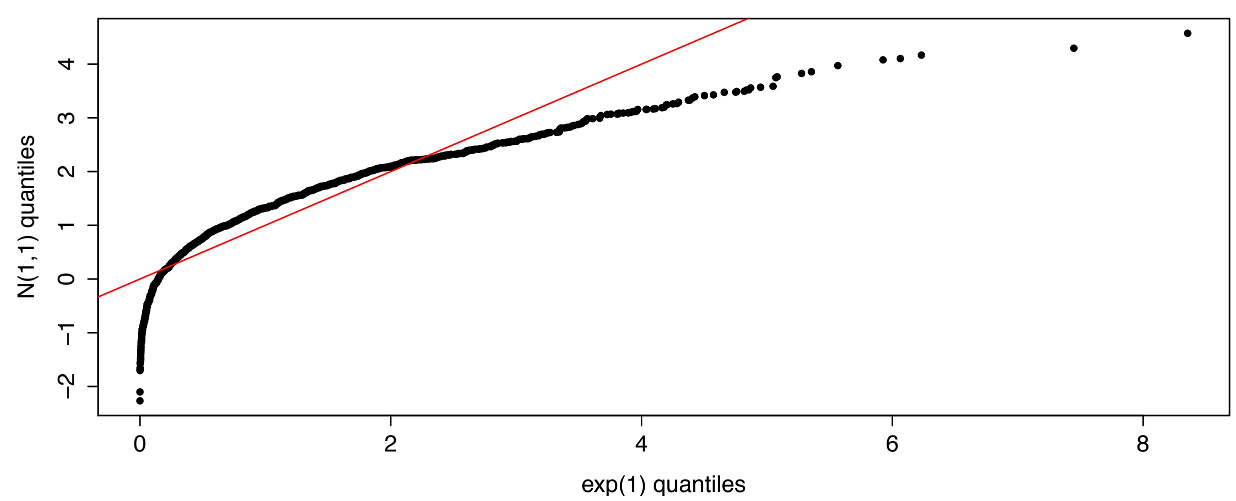

N(1,2) quantiles
−4
−2
0
2
4
6
−2
−1
0
1
2
3
4
N(1,1) quantiles



Histogram of a
Frequency
0
50
100
150
200
−2
0
2
4
6
8
a



Histogram of b
Frequency
0
20
40
60
80
−2
0
2
4
6
8
b



N(1,1) quantiles
−2
−1
0
1
2
3
4
0
2
4
6
8
exp(1) quantiles

But R has many built-in functions

binom

geom

pois

unif

exp

chisq

gamma

norm

t

...

Exponential task 1

- Generate 1000 Exp(λ) rvs. conditional on them each being greater than y, for some y (Try λ=1, y=1, say). Let's call those r.v.s X.
- Plot a histogram showing the distribution of (x-y), for y=1 and λ=1, and compare it to 1000 Exp(1) rvs. [or superimpose the exponential density function using the command curve(λ*exp(-λ*x)]
- How do we generate exponential rvs conditional on them being greater than y?

Simple rejection method

- To simulate 1000 exponential r.v.s:

  $u \sim U[0,1] \sim F(x)$. So $x \sim F^{-1}(u)$.

  Set $u=F(x)=1-\exp(-\lambda x)$

  $\exp(-\lambda x)=1-u$

  $x=(-1/\lambda)\log(1-u)=F^{-1}(u)$ [or, $x=(-1/\lambda)\log(u)$]

```
u<-runif(1000,0,1)
ExpRVs<- (-1/lambda)*log(u)
```

- To generate $X \sim \exp(\lambda)$, conditional on $(X>y)$:

```
x <-0
while (x<y){
   # Generate x~exp(λ)
   }
```

The x-value that results has the correct distribution. So repeat that process 1000 times. We'll return to rejection sampling later in the course.

Pseudocode (Week2-ConditionedExponential repo on Gthub)

```
set.seed(99999)
# repeat the following until you have 1000 conditioned exponential rvs.
u<-runif(1,0,1)
y<-1    # suppose we want to condition on the rv being bigger than 1
lambda<-2   # suppose we want exponentials with parameter 2
ConditionedExpRV<- (-1/lambda)*log(u)
while (ConditionedExpRV < y){
  u<-runif(1,0,1)
  ConditionedExpRV<- (-1/lambda)*log(u)
}
#Store the value of ConditionedExpRV
```

Exponential: Memoryless property

- Memoryless property:
  - If X is Exp(λ), then $f(x+y|X>y)=f(x)$ (i.e., x-y is still Exp(λ)).
  -

Exponential task 2: Waiting for a bus

- Suppose times between bus arrivals are distributed as T~exp(1).

1. Suppose we arrive at a bus-stop at some fixed time during the day (say after 10 hours). How long, on average, do we have to wait for a bus? [What if we arrive at a random time each day?]
2. If we get off one bus and wait for the next one to arrive on the same route, how long, on average, do we have to wait?
3. How long on average was the time between the arrival of the bus we caught and the one before it.
4. What is the expected time between any two buses?

Note: the mean of an exp(λ) r.v. is 1/λ.

See 'Week2-BusWaitingTimesExercise on Github

How big should my dam be?

A dam with volume V, releases water at a constant rate of L million liters per day.

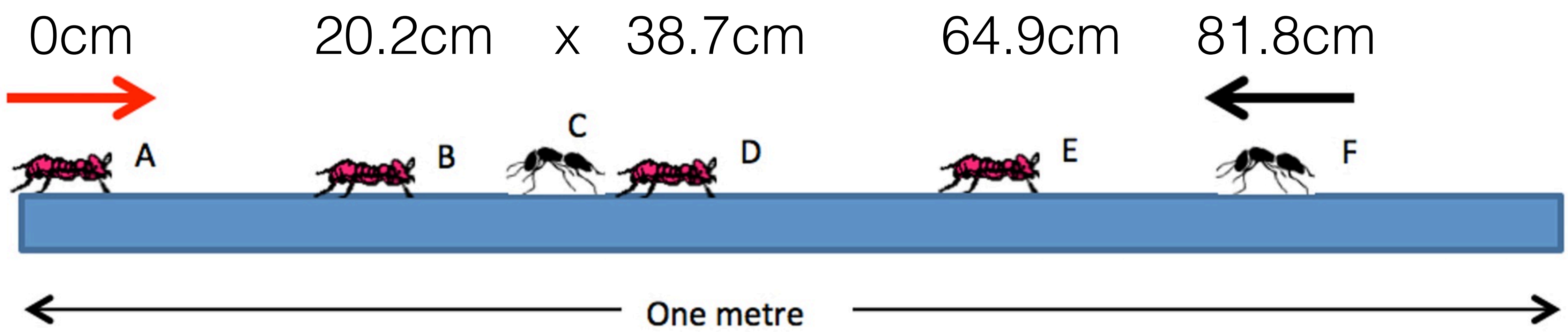
If the amount of rainfall each day (in millions of liters) is distributed as a N(5,10) r.v., what size dam do we need to make sure the probability that the dam will overflow at some point within the next year is less than 20%?

Is a theoretical answer possible?

Partial pseudocode on Blackboard as DamSimulator1_Pseudocode.R

.

Ants on a stick



- Ants walk at 1cm/second. When they meet, each ant turns around and walks in the other direction. When they reach the end of the stick they fall off.
  - How many seconds until the last ant falls off?
  - Which ant is the last to fall off the stick?

END