Randomized Algorithms

Worked Problems

Solutions

Student Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Professor: Nik Bear Brown

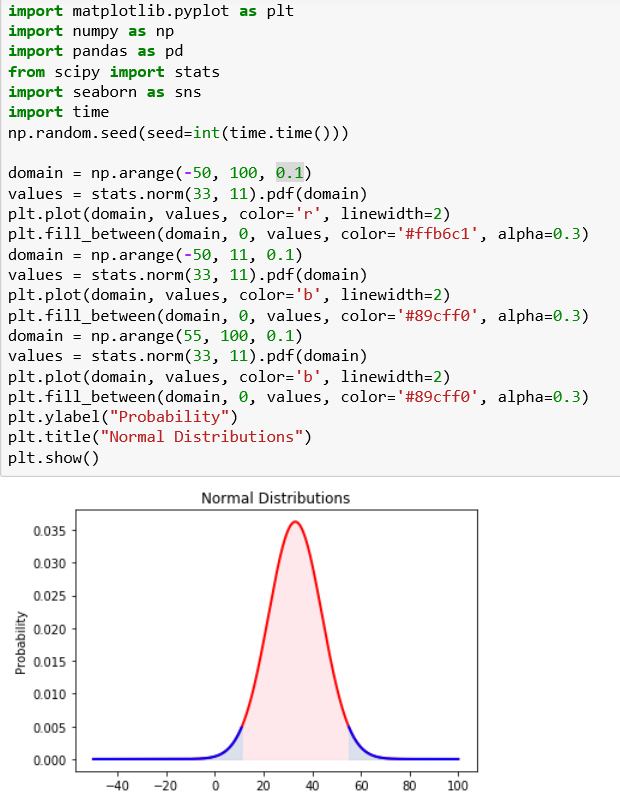
Q Given a normal distribution with a mean to 33, a standard deviation of 11, and the sample size to 100. What is the probability of finding a value:

1. less than 11 (2 points)
2. greater than 55 (2 points)
3. less than 11 or greater than 55 (1 point) Show the calculation as done by hand.

Because of normal distribution, we can directly calculate z on the basis of

1. When value is less than 11, z = (11 – 33) / (11 / = -20, P1(X < 11) = P1’( Z < -20) = 2.7536241186061556e-89
2. When value is greater than 55, U = (55 - 33) / (11 / = 2, P2(X > 55) = P2’(Z > 20) = 2.7536241186061556e-89
3. When value is less than 11 or greater than 55, means P3(X <11 or X > 55) = P3’(Z < -20 or Z > 20) = P1+ P2 = 5.507248237212311e-89

Qp Write python code to plot and calculate the above



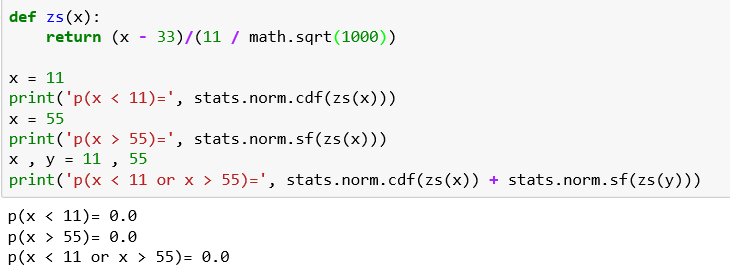
Q Given a normal distribution with a mean to 33, a standard deviation of 11, and the sample size to 1000. What is the probability of finding a value:

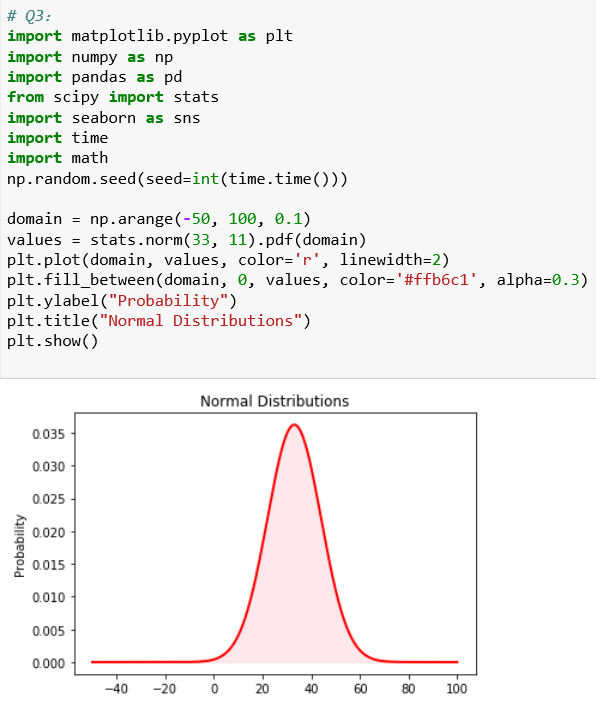
1. less than 11 (2 points)
2. greater than 55 (2 points)
3. less than 11 or greater than 55 (1 point) Show the calculation as done by hand.

Because of normal distribution, we can directly calculate z on the basis of

1. When value is less than 11, z = (11 – 33) / (11 / = -20, P1(X < 11) = P1( Z < -20) = 0
2. When value is greater than 55, U = (55 - 33) / (11 / = 20, P2(X > 55) = P2(Z > 20) = 0
3. When value is less than 11 or greater than 55, means P3(X <11 or X > 55) = P3(Z < -20 or Z > 20) = P1+ P2 = 0

Qp Write python code to plot and calculate the above





Q A company has placed an order for 5,000 laptops with a supplier on the condition that no more than 1% of the devices will be defective. To check the shipment, the company tests a random sample of 100 laptops and finds that 2 are defective.

Does this provide sufficient evidence to indicate that the proportion of defective can laptops in the shipment exceeds 1%? Explicitly state your null and alternative hypothesis.

We use μto denote the proportion of defective can laptops in the shipment

H0 : μ=> 1%

H1 : μ< 1%

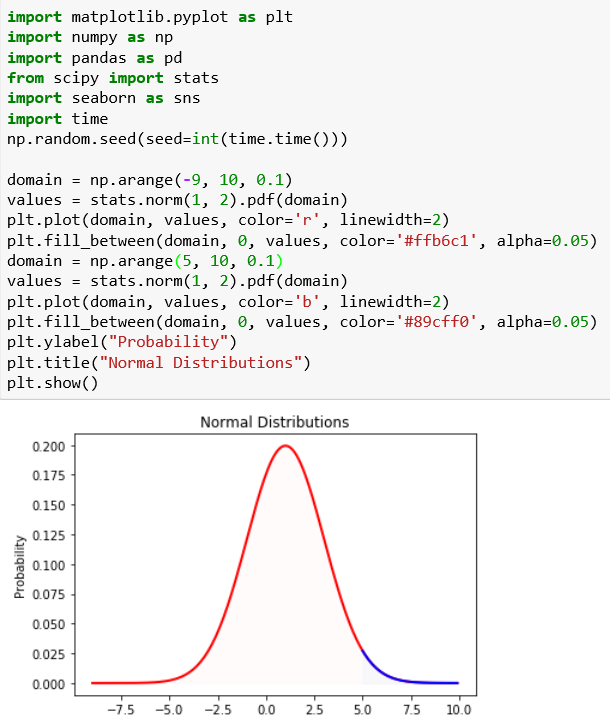
we give appropriate significant level α= 0.05

Because the sample fits in with the normal distribution, in the beginning, we can suppose the population variance equals 2. According to t-test formula,

= (2 – 1) / (2 / = 5 . From t-table, we can find that when t => 2.63, the proportion equals 0.05. So the rejection region {t | t > 2.63}

According to the data above, we conclude that the test cannot provide sufficient evidence to indicate that the proportion of defective can laptops in the shipment exceeds 1%.

Qp W Write python code to plot and calculate the above



Q An ultra-marathon runner ran 103 miles per week as reported by runners world. A random sample of 500 ultra-marathon runners had a mean of 101 miles per week ran when asked.

Let m denote mean distance for all ultra-marathon runners. A (3 Points).Perform the hypothesis test

Ho: m=103 miles per week ran Ha: m ≠103 miles per week ran

at the 5% significance level. Assume the standard deviation is 60 miles.

t = (103 - 101) / (60 / ) = 0.745, p = 0.23 > 5% significance level, so I do not have strong evidence against the null hypothesis

B (2 Points ). Find a 95% confidence interval for m.

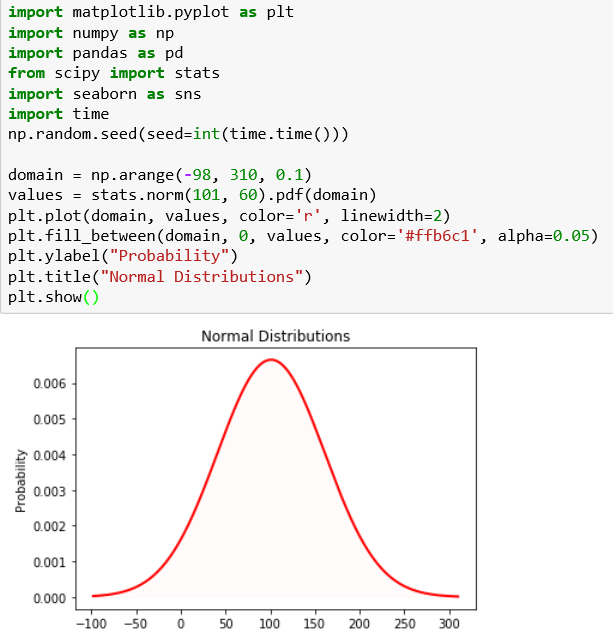
If we want to get a 95% confidence level, we need to make the p-value equals (1 – 95%) / 2 = 0.025. In other words, we should make sure that t equals 1.96

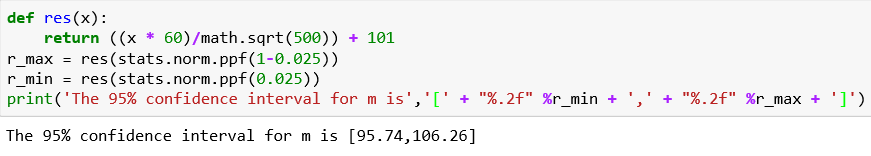
t = (m - 101) / (60 / ) = -1.96 , m = 96 or 106

We can use process to help us calculate accurate data: m ~ [95.74,106.26]

Show the calculation as done by hand.

Qp Write python code to plot and conduct a hypothesis test on the above





Q What is the probability of getting exactly 2 heads after flipping three coins?

Solution:

If one is head and 0 tails:

0 0 0

0 0 1

0 1 0

0 1 1

1 0 0

1 0 1

1 1 0

1 1 1

There are three events in getting exactly 2 heads

{0 1 1, 1 0 1, 1 1 0}

So 3/8 is the probability of getting exactly 2 heads after flipping three coins.

We can also the formula for a discrete random variable based on a binomial distribution:

****

X = 1, with probability 

Note: 3 choose 2 is 3. The 3/8 comes from 3\* 1/23

Q Consider a six-sided die that gets a 1 with probability p = 1/6. How confident are you that you can get a 1 after rolling the die 3 times?

Solution:

Because these are independent events (the roll of one die doesn’t affect another) we have p = 1/6 chance on each of the die throws. You want probability of at least a 1 in 3 rolls that means total probability (1 - none of the dice has 1). Hence the probability comes out to be 1−(5/6)3=1- 0.58=0.42

However if we want exactly one success (a roll of 1) in three tries this is just the binomial expansion. If we run ntrials, where the probability of success for each single trial is p, what is the probability of exactly k successes?

     ……. 

k slots where prob. success is , n-k slots where prob. failure is  ****

Thus, the probability of obtaining a specific configuration as denoted above is pk(1-p)n-k. From here, we must ask ourselves, how many configurations lead to exactly k successes. The answer to this question is simply, "the number of ways to choose k slots out of the n slots above. This is . Thus, we must add pk(1-p)n-k with itself exactly  times.

This leads to the formula:

****

X = 1, with probability 

Note: 3 choose 1 is 3

25/72 = 0.3472 see WolframAlpha - (3 choose 1)\* 1/6\*(5/6)^2 <http://po.st/ZztZy1>

Q Consider a coin that comes up heads 9 times out of ten flips. How confident are you that the coin is fair?

Solution:

Use this formula, using p=0.9 and n = 10 and solve for ε.

http://upload.wikimedia.org/math/e/0/2/e0221b0cee97e827245e1afae9f02ec2.png

Q

Does the payoff matrix below have a Nash equilibria? Why or why not?

|  |  |  |
| --- | --- | --- |
|  | Cooperate | Defect |
| Cooperate | 1,-1 | -1,1 |
| Defect | -1,1 | 1,-1 |

Solution:

No.

|  |  |  |
| --- | --- | --- |
|  | Cooperate | Defect |
| Cooperate | 1,-1 no -1->1 (player 2 defect) | -1,1 no -1->1 |
| Defect | -1,1 no -1->1 | 1,-1 no -1->1 |

Q

Does the payoff matrix below have a Nash equilibria? Why or why not?

|  |  |  |
| --- | --- | --- |
|  | Cooperate | Defect |
| Cooperate | 1,1 | 1,1 |
| Defect | 1,1 | 1,1 |

Solution:

Yes, any point is an equilibrium. The payoff is the same at any point so neither party will ever want to change from any point, making it a global equilibrium.

Q

Will the State Flipping Algorithm used for Hopfield networks always converge to an optimal solution? Give a proof that it will or it won’t.

Solution:

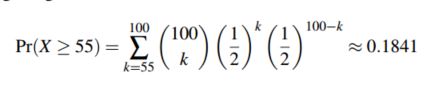
The State Flipping Algorithm used for Hopfield networks will NOT always converge to an optimal solution. This can be shown with any of many counter-examples. For example, the G be a graph consisting of a cycle of length four: there are nodes v1, v2, v3, v4 and edges (v1, v2), (v2, v3), (v3, v4),( v4, v1). Then if we start the state-flipping algorithm in a configuration where the nodes v1 and v2, have state +1 and the nodes v3 and v4 have state -1 then no improving move is possible.

Q: Determine the probability of obtaining 55 or more heads when flipping a fair coin 100 times by an explicit calculation, and compare this with the Chernoff bound.

Do the same for 550 or more heads in 1000 flips.

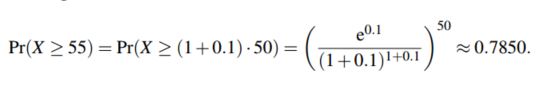
Solution:

Let X ∼ Bin(100,1/2) be the number of heads for 100 flips. The expectation of X is then µ = E[X] = 50. The exact probability of getting 55 or more heads is,

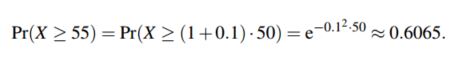


We can obtain several different Chernoff bounds for the same distribution.

Using the first bound,

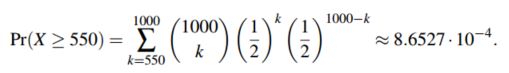


Alternatively,

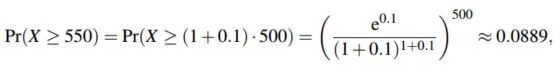


Comparing to the exact value we can observe that the first bound is about 4.3 times larger and the second bound is about 3.3 times larger.

Similarly, for 1000 flips, the number of heads is X ∼ Bin(1000,1/2), and the exact probability is



The first bound gives



and the second gives

6

This time the first bound is about 103 times larger than the exact value and the second bound is about 7.8 times larger. Thus, the relative error of the bounds seems grow when the number of flips increases.

Q Suppose that we have two identical boxes: box 1 and box 2. Box 1 contains 5 red balls and 3 blue balls. Box 2 contains 2 red balls and 4 blue balls. A box is selected at random and exactly one ball is drawn from the box.

a. What is the probability that the ball is blue?

b. Given that the selected ball is blue, what’s the probability that it came from box 2?

Solution:

a.

Let,

B1 = Event that Box1 is selected

B1 = Event that Box1 is selected

R = event that a Red Ball is selected.

B = event that a Blue Ball is selected.

Therefore,   
P(B1) = ½

P(B2) = ½

P(B | B1) = 3/8

P(B | B2) = 4/6 = 2/3

Therefore,

P(B) = P(B1) P (B|B1) + P(B2) P (B|B2) = (9+16) / 48 = 25/48 = ½

b. Using Bayes’ theorem,

P(B2|B) = P(B2) P(B|B2) / (P(B1) P(B|B1) + P(B2) P(B|B2) ) = (1/3) / (25/48) = 16 / 25.

Q How many people must be gathered together in a room, before you can be certain that there is a greater than 50/50 chance that at least two of them have the same birthday?

Solution:

Only twenty-three people need be in the room, a surprisingly small number.

The probability that there will not be two matching birthdays is then, ignoring leap years, 365x364x363x...x343/365 over 23 which is approximately 0.493.

This is less than half, and therefore the probability that a pair occurs is greater than 50-50. With as few as 23 people in the room the chances are better than 50-50 that a pair will have birthdays on the same day or on consecutive days.

Q. Three rats are sitting at the three corners of an equilateral triangle. Each rat starts randomly picks a direction and starts to move along the edge of the triangle. What is the probability that none of the rats collide?

Solution:

The ants can only avoid a collision if they all decide to move in the same direction (either clockwise or anti-clockwise). If the ants do not pick the same direction, there will definitely be a collision. Each ant has the option to either move clockwise or anti-clockwise. There is a one in two chance that an ant decides to pick a particular direction. Using simple probability calculations, we can determine the probability of no collision.

Let, n = probability of no collision   
n = x (Rats move clockwise) + y (Rats move in counter-clockwise)   
n = (.50 x .50 x .50) + (.50 x .50 x .50)   
n = (.125) + (.125)   
n = .25 or 25%

Q A monkey at a typewriter types each of the 26 letters of the alphabet exactly once, the order being random. What is the probability that the word Miley appears somewhere in the string of letters?

Solution:

Miley (a 6-letter string) could appear in 21 different spots in the 26-letter string (e.g., spot 1: Mileycxs…; spot 2: cMileycksd…; spot 3: ckMileydsfd…; etc. up to: spot 21: qwryuiopsdfghjkzxcvbnMiley

The remaining 20 letters can be arranged in any sequence (there are 20! ways to arrange the other 20 letters in the remaining 20 spots).

Therefore, there are 21*x*20! = 21! ways to get Miley.

There are 26 letters to fill 26 spots, so there are 26! total arrangements of the letters.

∴P(getting Miley) =  very small!

Q An elevator containing five people can stop at any of seven floors. What is the probability that no two people get off at the same floor? Assume that the occupants act independently and that all floors are equally likely for each occupant.

Solution:

Denominator (with replacement, since everyone could get off at the same floor): 

Numerator (without replacement): 

∴P(no two people on the same floor) =  15% chance

Q A card is drawn at random from a deck of 52 playing cards. What is the probability that it is a king or a heart?

Solution:



Q Describe the Metropolis algorithm. How does it differ from gradient descent?

Solution:

Metropolis algorithm

Globally biased toward "downhill" steps, but occasionally makes "uphill" steps to break out of local minima.

**Metropolis algorithm (symmetric proposal distribution)**

Let *f(x)* be a function that is proportional to the desired probability distribution *P(x)*.

1. Initialization: Choose an arbitrary point *x0* to be the first sample, and choose an arbitrary probability density Q(x|y) which suggests a candidate for the next sample value *x*, given the previous sample value *y*. For the Metropolis algorithm, *Q*must be symmetric; in other words, it must satisfy Q(x|y) = Q(y|x). A usual choice is to let Q(x|y) be a Gaussian distribution centered at y, so that points closer to y are more likely to be visited next—making the sequence of samples into a random walk. The function Q is referred to as the proposal *density* or *jumping distribution*.
2. For each iteration *t*:
   * Generate a candidate *x'* for the next sample by picking from the distribution Q(x'|x_t).
   * Calculate the *acceptance ratio* α = *f(x')/f(xt)*, which will be used to decide whether to accept or reject the candidate. Because *f* is proportional to the density of *P*, we have that *α = f(x')/f(xt)* = *P(x')/P(xt)*.
   * If α ≥ 1, then the candidate is more likely than *xt*; automatically accept the candidate by setting *xt+1 = x'*. Otherwise, accept the candidate with probability α; if the candidate is rejected, set *xt+1 = xt*, instead.

Gradient descent

Let S denote current solution. If there is a neighbor S' of S with strictly lower cost, replace S with the neighbor whose cost is as small as possible. Otherwise, terminate the algorithm.

Gradient descent can’t occasionally make "uphill" steps. Gradient descent is based on the observation that if the multivariable function F(\mathbf{x}) is defined and differentiable in a neighborhood of a point \mathbf{a}, then F(\mathbf{x}) decreases *fastest* if one goes from \mathbf{a} in the direction of the negative gradient of F at \mathbf{a}, -\nabla F(\mathbf{a}). It follows that, if  \mathbf{b} = \mathbf{a}-\gamma\nabla F(\mathbf{a}) for \gamma small enough, then F(\mathbf{a})\geq F(\mathbf{b})

Q Does the payoff matrix below have any Nash equilibria? Why or why not?

|  |  |  |
| --- | --- | --- |
|  | Cooperate | Defect |
| Cooperate | 2,2 | 0,1 |
| Defect | 1,0 | 1,1 |

Solution:

Yes (Cooperate, Cooperate) (2,2) and (Defect, Defect) (1,1). Neither player gains from switching. Note in (Defect, Defect) (1,1) we don’t gain but don’t lose either 1 to 1.

Q In a room of 23 people, what is the probability that someone has the same birthday as you?

Solution:

The probability *q*(*n*) that someone in a room of *n* other people has the same birthday as a particular person (for example, you), is given by

 q(n) = 1 - \left( \frac{365-1}{365} \right)^n 

and for general *d* by

 q(n;d) = 1 - \left( \frac{d-1}{d} \right)^n. 

In the standard case of *d* = 365 substituting *n* = 23 gives about 6.1%

See <http://en.wikipedia.org/wiki/Birthday_problem#Same_birthday_as_you>

Q A coin seems biased, say coming up heads 75% of the time , how many times to you have to flip it to determine its bias with a 95% confidence (ε = .05) ?

Solution:

 n \geq \frac{1}{(p -\frac{1}{2})^2} \ln \frac{1}{\sqrt{\varepsilon}}. http://www4b.wolframalpha.com/Calculate/MSP/MSP41381fcfg16214i39iee00005b25d6f08e840f2b?MSPStoreType=image/gif&s=63&w=171.&h=42.

About 24 times. See WolframAlpha - 1/(0.75 -0.5)^2 ln(1/squareroot(0.05)) <http://po.st/oZQRAb>

via @wolfram\_alpha

Q

Suppose S is a very large set of real numbers, and you would like to estimate its median value by sampling. (It is too expensive to sort S to calculate the exact median). You may assume all the numbers in S are distinct. Let n =|S|; a number x is an *ε-approximate median* if at least (0.5 – ε)n numbers in S are less than x and (0.5 – ε)n numbers in S are greater than x.

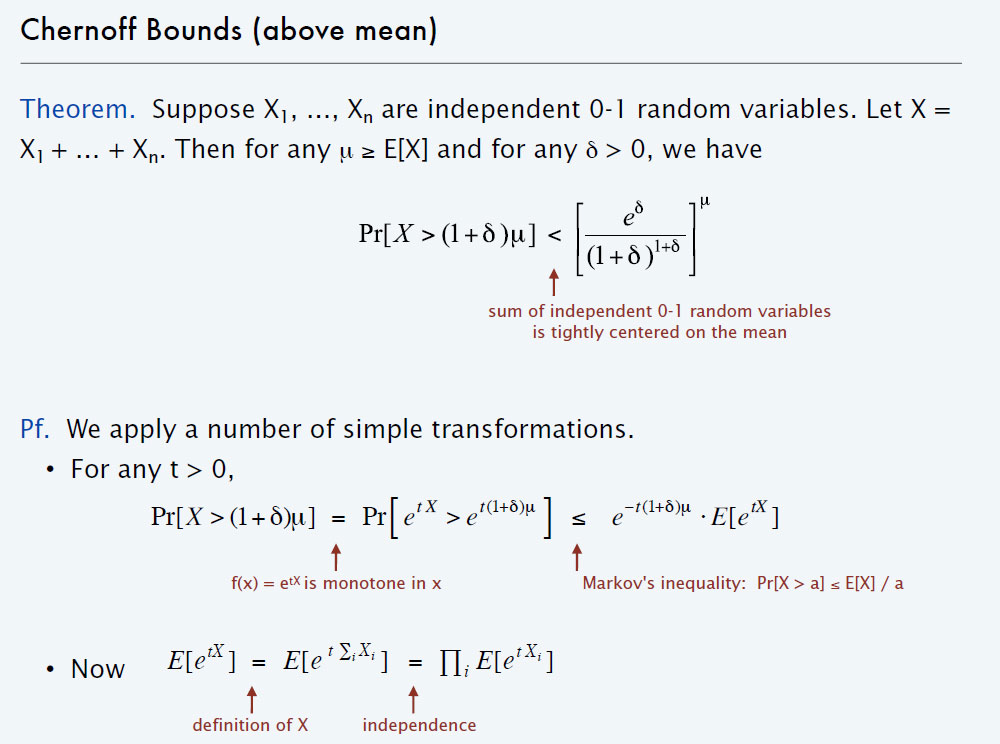
Suppose you sample a subset uniformly at random (with replacement) and use that to estimate your median. Use a Chernoff-bound to calculate your confidence in your approximate median estimate. You can choose your confidence level (typical is 90%, 95% or 99% confidence) the distribution of the numbers and your sample size.

Solution:

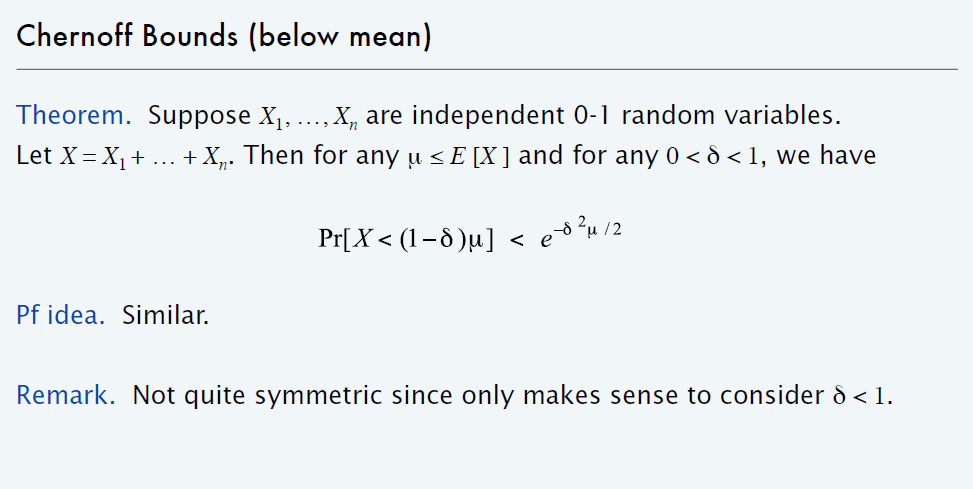
Note it’s OK just to pick a confidence level, and your sample size and use a Chernoff bound above and below the mean. To understand these Chernoff bounds you need to understand there are two sources of uncertainty in this process. There’s the probability that you obtain a “good” estimate. And there’s the extent to which your estimate is “good”. If you look at professional poll results that you’ll find that they use phrases like, “This poll is accurate to within ±5%, 19 times out of 20.” The ±5% is the range of the estimate and the 19 times out of 20 (95%) is the confidence in the estimate. Think of the coin flip bias check examples. If we wanted to check a 90% biased coin our range is very large ±40% so we need very few flips to get an estimate of high confidence. If we wanted to check whether a coin was slightly biased, say ±2%, we would need a lot of flips to estimate this with high confidence. So you will receive full credit for just deciding a confidence level and an estimate range to determine the sample size by plugging values into an appropriate Chernoff bound equation.

See Chernoff bound equations below:

Chernoff bound (above mean)



Chernoff bound (below mean)



The answer below is Kleinberg and Tardos preferred approach to this problem. This is also accepted but not expected.

Suppose we want 95% confidence (an *ε-approximate median* of .05). Imagine dividing the set S into 20 quantiles Q1 .. Q20 where Qi, consists of all elements that have at least. .05(i — 1)*n* elements less than them, and at. least .05(20 — i)*n* elements greater than them. Choosing the sample S' is like throwing a set of numbers at random into bins labeled with Q1 .. Q20.

Suppose we choose |S’| = 40,000 and sample with replacement. Consider the event E that S' intersection with Qi is between 1900 and 2100 for each i. If E occurs, then the first nine quantiles contain at most 19, 900 elements of S', and the last nine quantiles do as well. Hence the median of S' will belong to Q10 union with Q11 and thus will be a (.05)-approximate median of S.

The probability that a given Qi, contains more than 2100 elements can be computed using the Chernoff bound (above mean) with µ = 2000 and δ = .05; it is less than

Pr[X>2100]

Pr[X>(1+ δ) µ] < [*e*.δ /(1+ δ) (1+ δ)] µ

Pr[X>(1+ δ) µ] < [*e*.05/(1.05)( 1.05) ]2000

(i.e. [*e*^.05/(1.05)^( 1.05) ]^2000

Pr[X>(1+ δ) µ] is Pr[X>(1.05)2000] is Pr[X>2100]

The probability that a given Qi, contains fewer than 1900 elements can be computed using the Chernoff bound (below mean), with µ = 2000 and δ = .05; it is less than

Pr[X<1900]

Pr[X<(1- δ) µ] < [*e*-δ^2 µ/2 /(1+ δ) (1+ δ)] µ

*e*-(.5)( .05)( .05)200  (i.e. [*e*^-(.5)( .05)( .05)2000)

We then apply the Union Bound over the 20 choices of I, to get the probability that E does not occur.