Lin Chen
2/15/15

Solutions for the screening test from Qi Shi

Math:

Given a stick, if randomly cut into N pieces, the probability that the N pieces can form an N sided polygon?

Solutions:

If we randomly cut a stick into N pieces and each piece length is X_i, i=1,...,N. In this way, we cut the stick with (N-1) cut points. Without loss generality, we assume the length of the stick equal to 1. So,

$$\sum_{i=1}^{N} X_i = 1 \tag{1.1}$$

The cut of a stick with (N-1) cut points is equivalent with the cut of a circle with N cut points, where the length of the stick is equal to the circumference of the circle. Suppose the stick is soft and can be transformed to a circle (like a rope). Now, the circumference of the circle is equal to 1. We randomly cut this circle into N parts with N cut points. The arc of each part is X_i, i=1,...,N. (Suppose each arc can be transformed in to a line again). These two ways of cut are equivalent to form a polygon.

Next, we sort these random pieces in order, $X^{(1)}$,....,$X^{(N)}$, where $X^{(1)}$ is the piece with the shortest length and $X^{(N)}$ is the longest one. (Since these X_i are continuous random variables, the probability that all pieces length are equal is zero). In order to construct a polygon, these pieces must satisfy,

$$X^{(N)} < \sum_{k=1}^{N-1} X^{(k)} \tag{1.2}$$

Thus, the probability that the N pieces can form an N sided polygon is

$$P(N \text{ pieces can form an N sided plygon}) = P(X^{(N)} < \sum_{k=1}^{N-1} X^{(k)}) = 1 - P(X^{(N)} > \sum_{k=1}^{N-1} X^{(k)}) \tag{1.3}$$

If we think the cut with the circle case, the probability that we cannot construct a N-sided polygon $P(X^{(N)} > \sum_{k=1}^{N-1} X^{(k)})$ is equivalent with the probability that N cut points (or points) are located within a semicircle, $P(\text{N random located points are located wit hin a semicircle})$. This means the longest piece is longer than the half circle, thus $X^{(N)} > \sum_{k=1}^{N-1} X^{(k)}$. Given a cut point, i, we define a event E_i as: Starting at point, i, the next (N-1) points are in the clockwise semicircle and $P(E_i) = 1/2^{N-1}$. Thus, $P(X^{(N)} > \sum_{k=1}^{N-1} X^{(k)}) = P(\bigcup_{i=1}^{N} E_i)$. Since the events E_i and E_j, i≠j, are mutually exclusive, (Starting the ith point and including all points within the clockwise semicircle clockwise will not include the case t hat starting the jth point and including all points within the clockwise semicircle where i≠j, since for the ith point case ,there must be a point whose clockwise distance for jth point are larger than the half circle)

$$P(X^{(N)} > \sum_{k=1}^{N-1} X^{(k)}) = P(\bigcup_{i=1}^{N} E_i) = \sum_{i=1}^{k} P(E_i) = N / 2^{N-1} \tag{1.4}$$

So, from the Equation (1.3), we can say

$$P(N \text{ pieces can form an N sided plygon}) = 1 - P(X^{(N)} > \sum_{k=1}^{N-1} X^{(k)}) = 1 - N / 2^{N-1}.$$

Math Plus:

Given two functions: f(x) = x^a; g(x) = (ln(x))^b, where x>0; both a > 0, b > 0. Please compare f(x) against g(x) for all different scenarios of (a, b) combinations.

Hint: a) basically, (a,b) can be any point in the first quadrant

b) try to consider the full spectrum as x varies in (0,∞)

c) if plot f(x) against g(x), how many times does f(x) intersect g(x)

Solutions:

First, we consider the case that $x \geq 1$. We apply a one-to-one mapping by setting $x = \exp(z/a)$. Since a > 0 and $x \geq 1$, the $z \in [0, +\infty)$. Thus,

$$\begin{aligned} f(x) = x^a &=> f(z) = \exp(z) \\ g(x) = (ln(x))^b &=> g(z) = (z/a)^b \end{aligned} \tag{2.1}$$

We know that for any pair (a,b), the $\exp(z)$ always increases faster than $(z/a)^b$ as z increases and both of f(z) and g(z) are monotonically increasing function. So, there exists a z_0, such that $\exp(z) > (z/a)^b$ for $z > z_0$. Also, we know that $f(z=0) > g(z=0)$. So, given a pair (a,b), if we can find that $g(z) > f(z)$ for $z \in [0, z_0)$, then we can claim that there are two intersection points for $g(z)$ and $f(z)$, when $z > 0$. Or equivalently, there are two intersection points for $g(z)$ and $f(z)$, when $x \geq 1$. If $g(z) = f(z)$ for $z \in [0, z_0)$, there is one intersection point. Otherwise, there is no intersection between these two functions. In order to compare f(z) and g(z), we take a log of the ratio between g(z) and f(z),

$$h(z) = \log\left(\frac{g(z)}{f(z)}\right) = \log\left(\frac{(z/a)^b}{\exp(z)}\right) \\ = b(\ln z - \ln a) - z \tag{2.2}$$

if there exists z' such that g(z')>f(z'), h(z')>0. We calculate first and second derivative of h(z),

$$h'(z) = b/z - 1 \\ h''(z) = -b/z^2 \tag{2.3}$$

Since b>0, h''(z)<0. These exists a maximum value for h(z) when h'(z)=0. From Equation (2.3) we get when z'=b , h(z) approach its maximum $h(z'=b) = b(\ln b - \ln a) - b = b\left(\ln\frac{b}{a}\right) - b$. If $h(z'=b) > 0$, we can get $\frac{b}{a} > e$. Here, $h(z'=b) > 0$ means that there exist some z>0 such that g(z)>f(z). According to the discussion above, we can claim that, given z>0, if $\frac{b}{a} > e$, there exists two intersection points between g(z) and f(z). If $\frac{b}{a} = e$, there exists one intersection points for g(z) and f(z). If $\frac{b}{a} < e$, there is no intersection point for g(z) and f(z).

Next, we consider the case that $0 < x < 1$. First, we set $w = 1/x$, where $w \in (1, +\infty)$ and f(x) and g(x) become,

$$f(w) = (1/w)^a \\ g(w) = (ln(1/w))^b \tag{2.4}$$

where w > 0. Similar to the first section, we take a one-to-one mapping. $w = \exp(u/a)$ where $u \in (0, +\infty)$. Thus

$$f(u) = \exp(-u) \\ g(u) = (-1)^b (u/a)^b \tag{2.5}$$

Since, f(u)>0 and so if $(-1)^b < 0$ or $(-1)^b$ is imaginary number, there will be no intersection points between f(u) and g(u). For the case when $(-1)^b$ is a positive number, there will be an interaction point between f(u) and g(u) for u>0 since f(u) is monotonically decreasing function ($\mathrm{f}(0) = 1$ and $\mathrm{f}(+\infty) = 0$) and g(u) is an monotonically increasing function due to a>0, b>0 ($g(0) = 1$ and $\mathrm{g}(+\infty) = +\infty$).

In summary,

Lin Chen
2/15/15

for a>0, b>0 and x>0,

i If $(-1)^b$ is a positive real number:

i(a) if $\frac{b}{a} > e$, there will be three intersection points between f(x) and g(x).

i(c) if $\frac{b}{a} = e$, there will be two intersection points between f(x) and g(x).

i(b) if $\frac{b}{a} < e$, there will be one intersection point between f(x) and g(x).

iI If $(-1)^b$ is not a positive real number:

iI(a) if $\frac{b}{a} > e$, there will be two intersection points between f(x) and g(x).

iI(b) if $\frac{b}{a} = e$, there will be one intersection point between f(x) and g(x).

iI(c) if $\frac{b}{a} < e$, there will be zero intersection point between f(x) and g(x).

Programming:

Implement a program to find out whether there exist M days within the last N (N>=M) trading days that the average closing price of these M days is at most P. Assume we have collected the history of the closing prices of the last N trading days for a stock.

Idea:

The idea is like this. When we load the stock price array size N, the sub-array size M and the price threshold p, we automatically generate a day index, a[], for the stock. a[0]=0 ,..., a[N]=N-1. Then, we sort the day index according to the corresponding stock price by merge sort method.

The example is like this, if we have 3 days stock prices as S=[8.1, 7.6, 6.2]. The corresponding day index will be a=[0,1,2]. Then, we sort the array according to the corresponding stock price in increasing order and save the result in the array, b. So, the b should be b=[2,1,0].

With the sorted day index array b, we can calculate the average stock price over the first M days. If this average is greater than p, then there do not exit such sub-array with size M that can give the average price lower than p.