

R 語言教學

鍾旻錡, 陳柏瑜

Statistics with Recitation

NTU Econ

2020.11.11

Outline

- 1 統計量
- 2 for loop
- 3 while loop
- 4 Monte Carlo Simulation

統計量

動差

- `mean(X)`
- `sd(X)`
- (Moments) `skewness(X)`
- (Moments) `kurtosis(X)`

其他常用統計量

- 相關係數
- 百分位數

查表

- 由 cdf 查百分位數
- 由百分位數查 cdf
- 由百分位數查 pdf (pmf)

Z Table

TA will teach you how to use the $N(0,1)$ table.



STANDARD NORMAL TABLE (Z)

Entries in the table give the area under the curve between the mean and z standard deviations above the mean. For example, for $z = 1.25$ the area under the curve between the mean (0) and z is 0.3944.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0190	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2487	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2969	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3513	0.3554	0.3577	0.3529	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4895	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

for loop

數列

- Summation

$$1 + 2 + \cdots + 100$$

- Fibonacci sequence

$$F_0 = 0, F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2}, \quad n = 2, 3, \cdots$$

練習

- $\sqrt{1} + \sqrt{2} + \cdots + \sqrt{100}$
 - 利用 `sqrt` 函數開根號

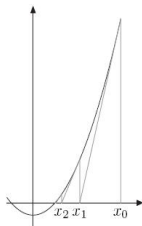
while loop

Geometric distribution

- 利用 Uniform 生成幾何分配
 - 由 Uniform 生成 Bernoulli 分配
 - 再由 Bernoulli 分配生成 Geometric 分配
- (補充) 大家可以自行嘗試將成功次數改成 2 次, 3 次或更多次, 即為負二項分配

Newton's method

- 說明



- 條件

- Function is smooth enough
- Good initial guess

Newton's method

- 找出多項式

$$f(x) = x^3 + 2x^2 - 7$$

之一根

- 求出

$$\sqrt[5]{10} = ?$$

猜拳

- 模擬 A 和 B 兩人猜拳
 - 利用 `sample` 函數使兩人隨機出拳
 - 利用 `if-else` 函數判斷勝負

練習

- 讓 A 和 B 連續猜拳, 先贏 5 次者勝, 看是 A 勝或 B 勝?
 - 建立變數儲存兩人勝利次數
 - 利用 `while` 迴圈使兩人不斷猜拳
 - 利用 `break` 終止迴圈

Monte Carlo Simulation

簡介

指使用亂數來解決很多計算問題的方法

- 所求解問題可以轉化為某種隨機分布的特徵數
- 問題本身具有內在的隨機性, 藉助電腦的運算能力可以直接類比這種隨機的過程

處理複雜計算

- 估計 π

$A = \text{number of } X^2 + Y^2 \leq 1$

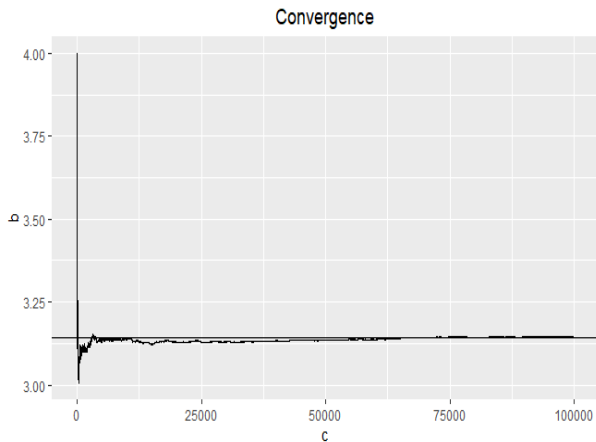
$B = \text{total points}$

由樣本類比母體, 因此

$$\frac{A}{B} = \frac{\pi}{4} \quad \Rightarrow \quad \pi = 4 \times \frac{A}{B}$$

處理複雜計算

- 收斂速度



練習

- 估計 $\sqrt{2}$

處理統計問題

- 計算幾何分配的平均值
- 畫出幾何分配的 pmf

處理統計問題

- (Thinking)

Suppose that we continually roll a die until the sum of all throws exceeds 100. What is the most likely value of this total when you stop?

感謝大家聆聽