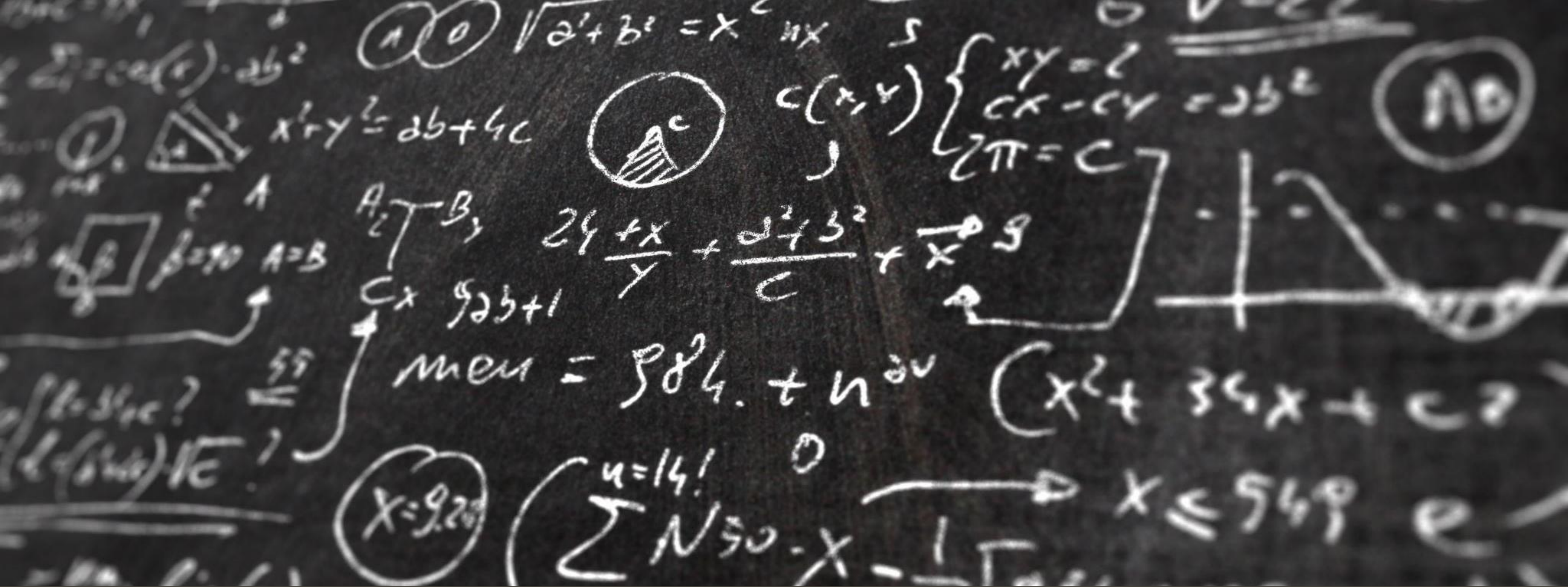


BASIC MATH & STATISTICS (PART II)

พัณณธัญญ์ วิจิตรวงศ์เจริญ

สารบัญ

- Joint distributions, covariance
- Normal Distribution
- Central limit theorem
- การประมาณค่า

JOINT PROBABILITY DISTRIBUTIONS

ความน่าจะเป็นของตัวแปรสุ่ม ตั้งแต่ 2 ตัวขึ้นไป

$g(x,y)$

คุณสมบัติ

มีค่ามากกว่าหรือเท่ากับ 0

รวมกัน = 1

EXPECTED VALUE OF THE RANDOM VARIABLE G(X,Y)
ค่าเฉลี่ยที่เป็นฟังก์ชันของตัวแปรสุ่มร่วมกัน 2 ตัว

Z=g(x,y)

$$E[g(X,Y)] = \sum_{y} \sum_{x} g(x,y).f(x.y)$$

$$\boxed{E(X,Y) = \sum_{y} \sum_{x} xy.f(x.y)}$$

EXPECTED VALUE OF THE RANDOM VARIABLE G(X,Y) ค่าเฉลี่ยที่เป็นฟังก์ชันของตัวแปรสุ่มร่วมกัน 2 ตัว

$$E\,[g(X,Y) \pm h(X,Y)] = E[g(X,Y)] \pm E[h(X,Y)]$$

$E\,(XY) = E(X).E(Y)$ เมื่อตัวแปรเป็นอิสระต่อกัน

COVARIANCE OF RANDOM VARIABLE X AND Y
ค่าความแปรปรวนร่วมของตัวแปรสุ่ม

$$\sigma_{xy} = \mathrm{E}[(X - \mu_x)(Y - \mu_y)$$
$$= \sum_x \sum_y (x - \mu_x)(y - \mu_y)\, f(x, y)$$

$$\sigma_{xy} = \mathrm{E(XY)} - \mu_x\, \mu_y$$

COVARIANCE OF RANDOM VARIABLE X AND Y
ค่าความแปรปรวนร่วมของตัวแปรสุ่ม

$$\sigma^2_{aX+b} = a^2\sigma^2_X$$

$$\sigma^2_{aX+bY} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY}$$

เมื่อตัวแปรเป็นอิสระต่อกัน

$$\sigma^2_{a_1x_1+a_2x_2+\ldots+a_nx_n} = a_1^2\sigma^2_{x_1} + a_2^2\sigma^2_{x_2} + \ldots + a_n^2\sigma^2_{x_n}$$

ตัวอย่าง

โรงงานแห่งหนึ่ง ทำการตัดกระดาษสี 3 สีคือ สีแดง ขาว และน้ำเงิน กำหนดให้ X,Y,Z เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยที่ X,Y และ Z เป็นความกว้างของแถบกระดาษสีแดง ขาว และน้ำเงินตามลำดับ กำหนดให้แถบกระดาษมีความยาวเท่ากัน สำหรับกระดาษทุกสี และให้

$\mu_X = \mu_Y = 4$ นิ้ว $\mu_Z = 6$ นิ้ว และ $\sigma_X = 0.5$ นิ้ว $\sigma_Y = 0.3$ นิ้ว $\sigma_Z = 0.4$ นิ้ว

W เป็นความกว้างของธงชาติไทย

W = 2X+2Y+Z

จงหาความกว้างเฉลี่ยของธงชาติไทย E(W) และความแปรปรวนของความกว้างนี้ Var (W)

ตัวอย่าง

W = 2X+2Y+Z

$\mu_X = \mu_Y = 4$ นิ้ว $\mu_Z = 6$ นิ้ว และ σ_X=0.5 นิ้ว $\sigma_Y = 0.3$ นิ้ว $\sigma_Z = 0.4$ นิ้ว

E(W) = 2E(X)+2E(Y)+E(Z)

= (2*4)+(2*4)+6 =22 นิ้ว

Var (W) $= 2^2\sigma_X^2 + 2^2\sigma_Y^2 + \sigma_Z^2$

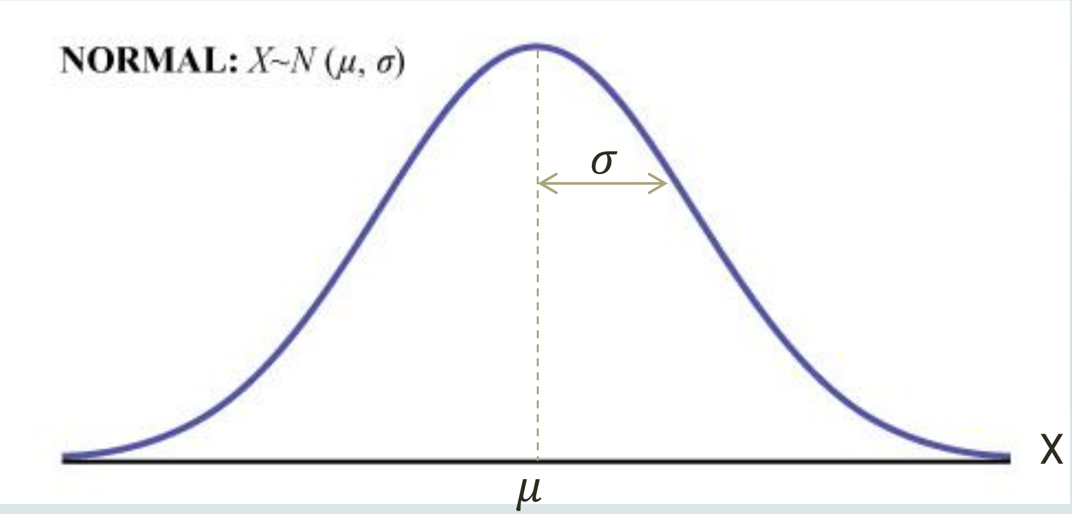
$= (4*0.5^2)+(4*0.3^2)+0.4^2 = 1.52$ นิ้ว

NORMAL DISTRIBUTION การแจกแจงปกติ

$$p(X \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right)$$

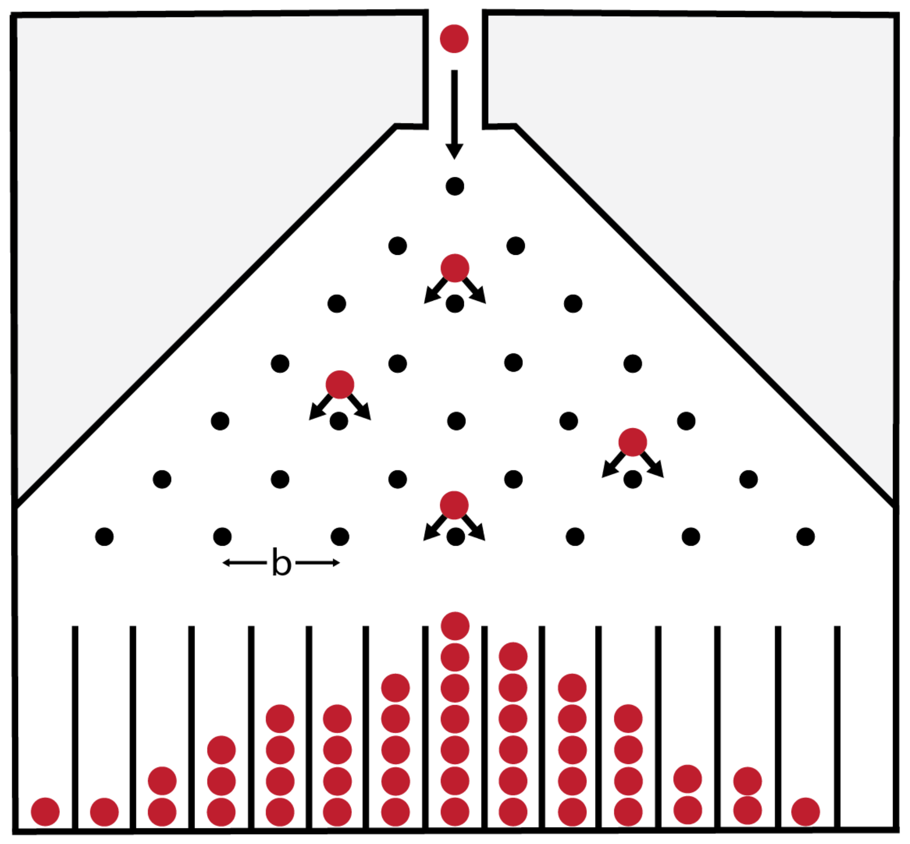
ตัวแปรสุ่ม X มีการแจกแจงเป็นรูประฆัง เรียกว่า ตัวแปรสุ่มปกติ

พารามิเตอร์ : μ , σ^2



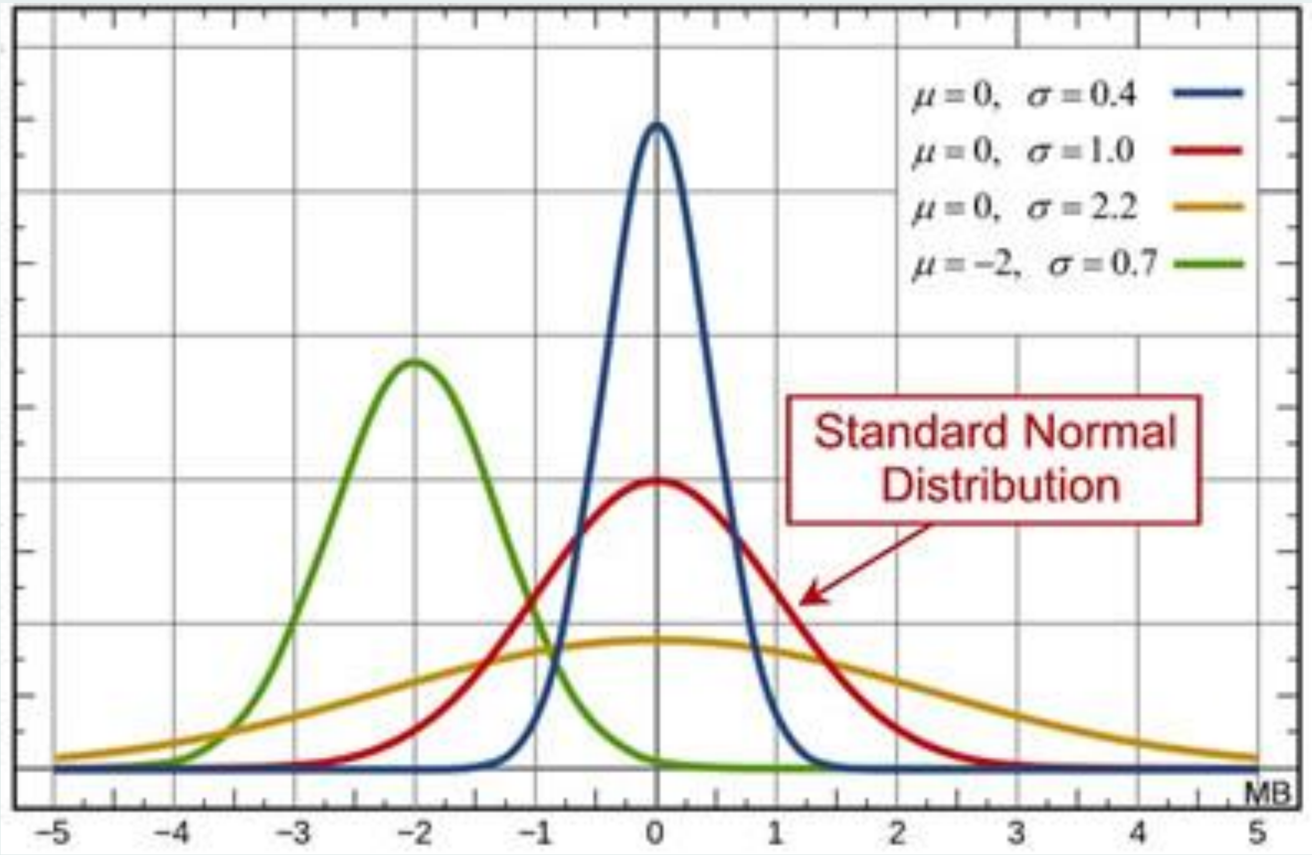
NORMAL DISTRIBUTION การแจกแจงปกติ

Galton Board

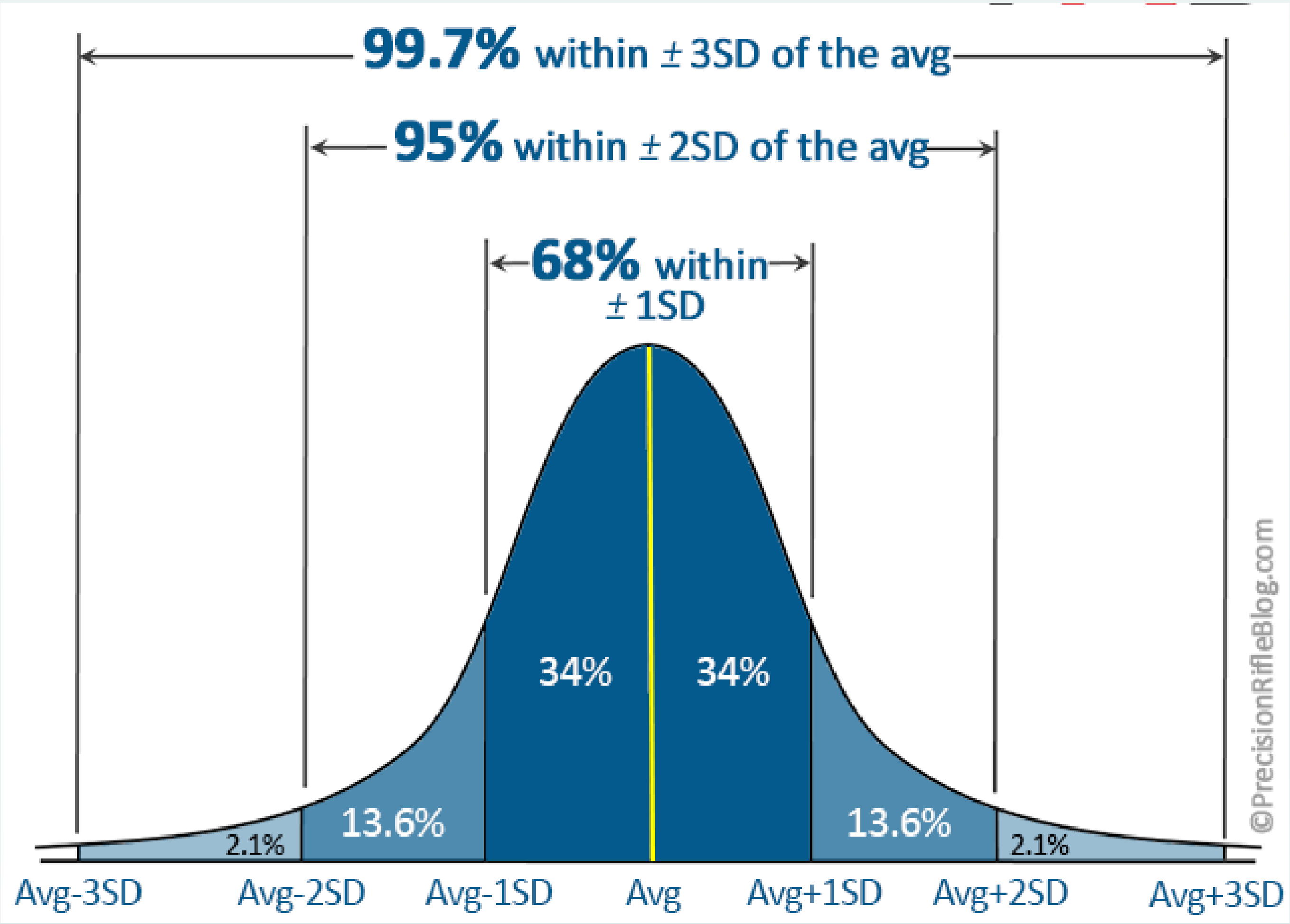


NORMAL DISTRIBUTION การแจกแจงปกติ

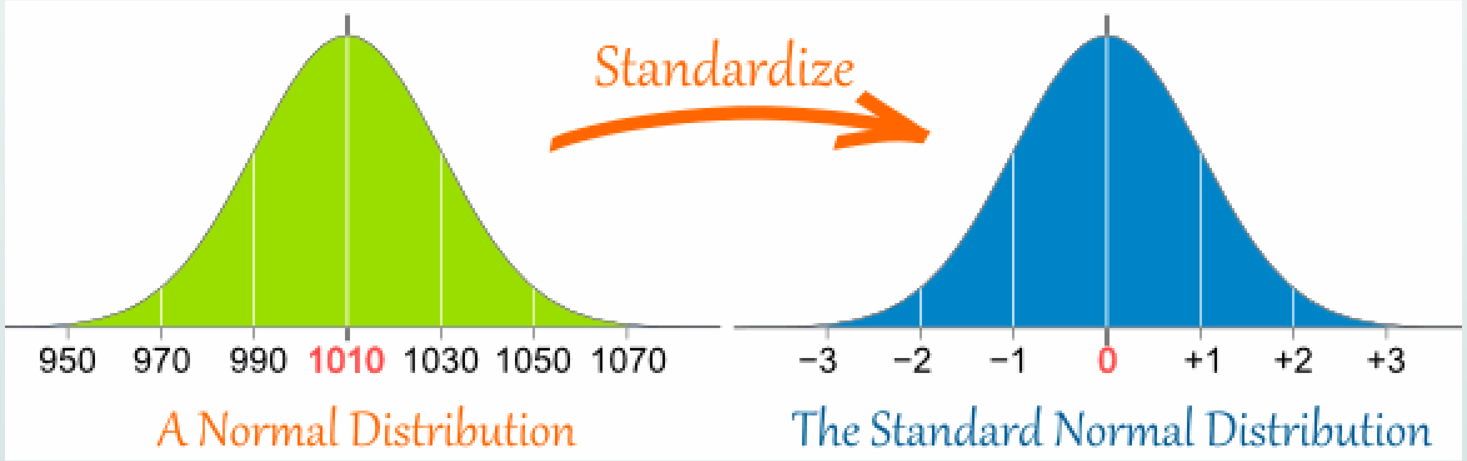
NORMAL DISTRIBUTION การแจกแจงปกติ



NORMAL DISTRIBUTION การแจกแจงปกติ



NORMAL DISTRIBUTION การแจกแจงปกติ

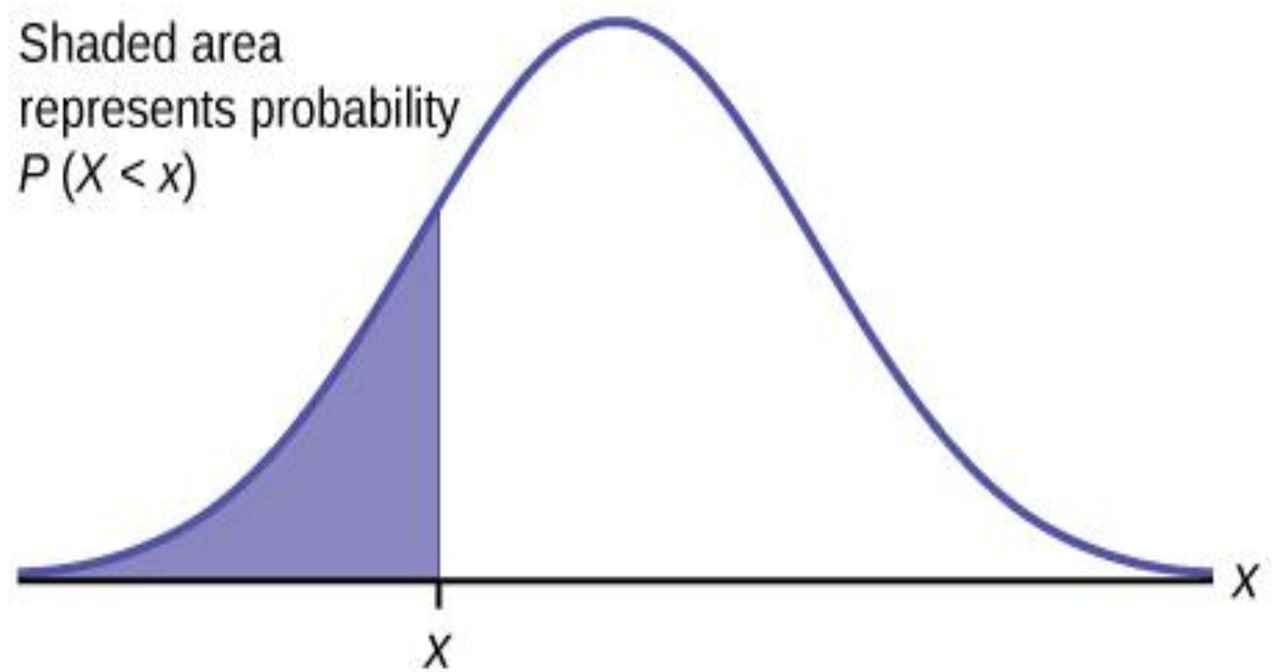


Std. normal distribution

$$\mu = 0$$
$$\sigma = 1$$

$$Z = \frac{X - \mu}{\sigma}$$

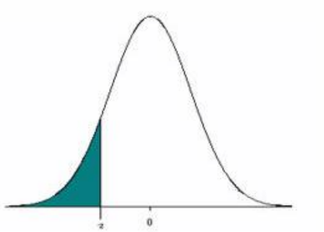
NORMAL DISTRIBUTION การแจกแจงปกติ



Standard Normal Distribution

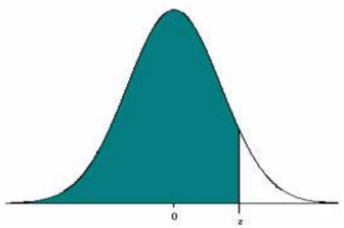
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Table of Standard Normal Probabilities for Negative Z-scores



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

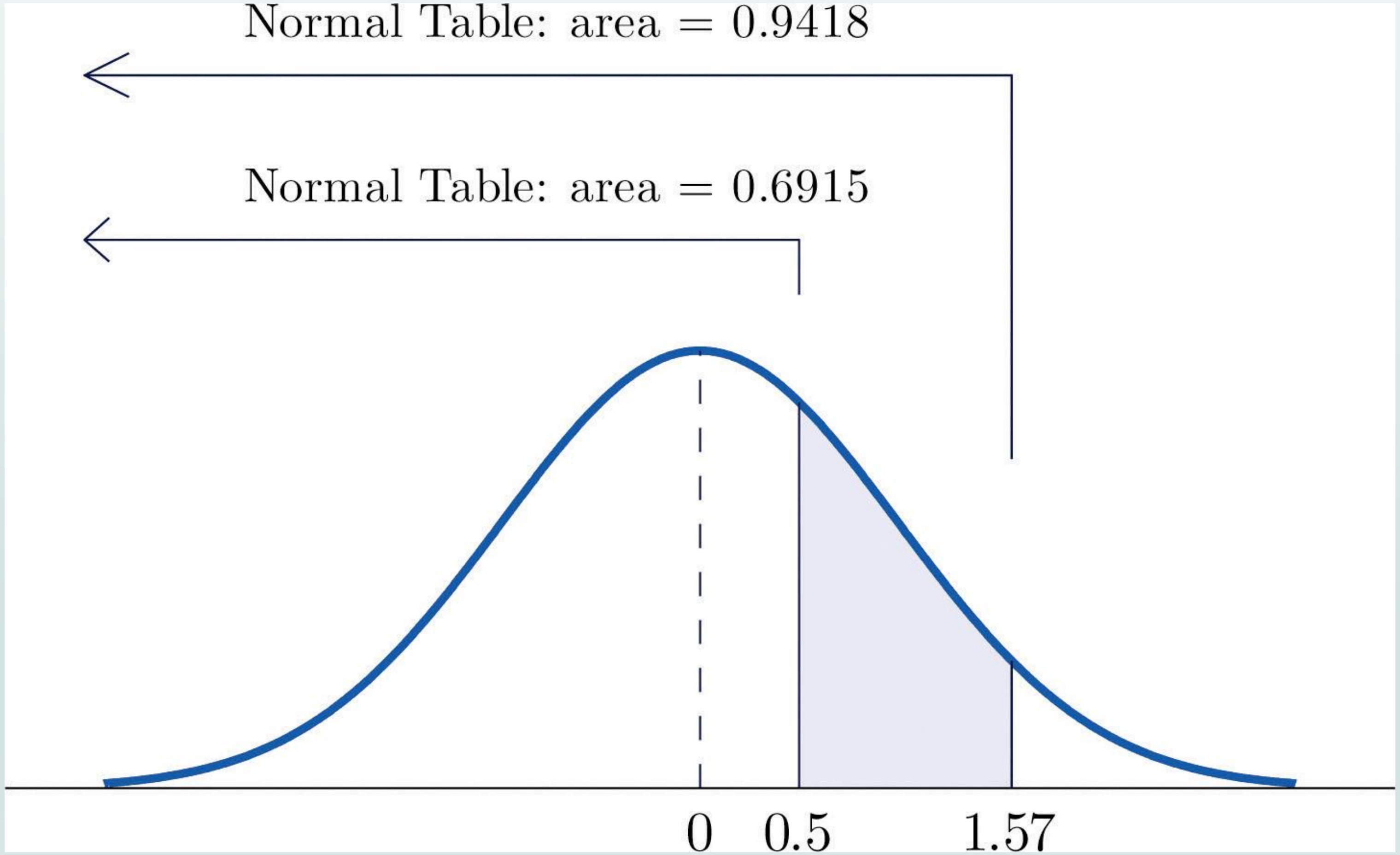
Table of Standard Normal Probabilities for Positive Z-scores



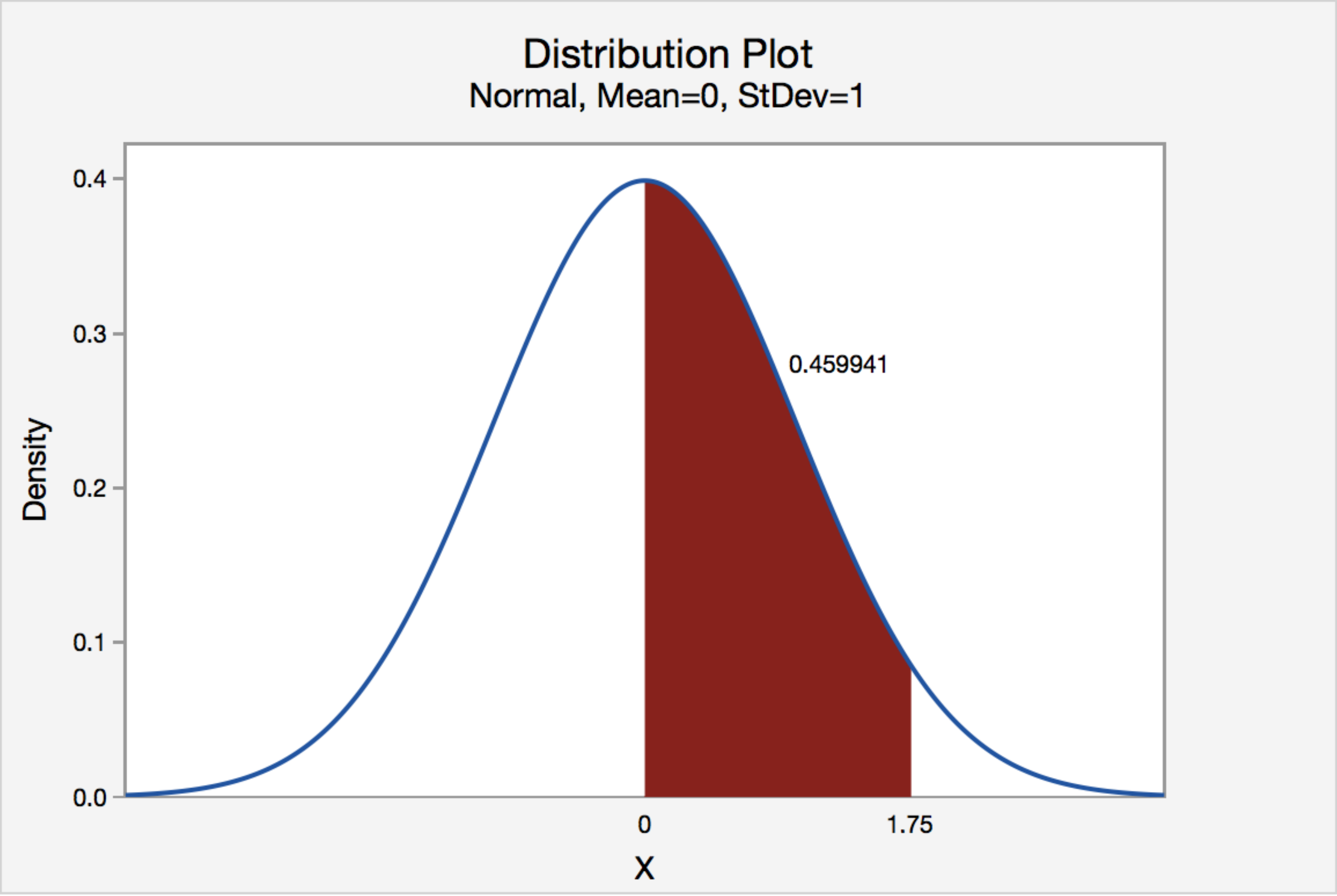
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Note that the probabilities given in this table represent the area to the LEFT of the z-score.
The area to the RIGHT of a z-score = 1 – the area to the LEFT of the z-score

NORMAL DISTRIBUTION การแจกแจงปกติ



NORMAL DISTRIBUTION การแจกแจงปกติ



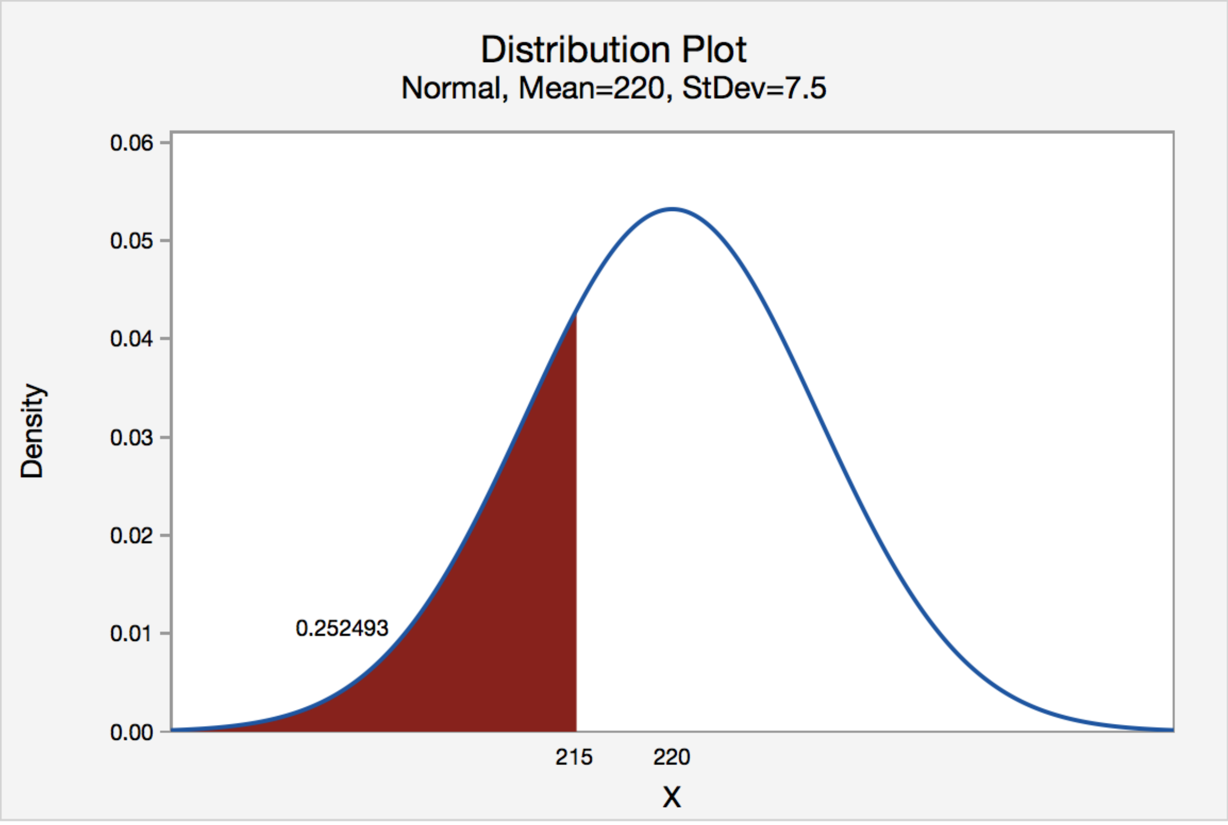
$$P\ (0 < x < 1.75)$$

= P(x<1.75) - P(x<0)

= 0.9599 – 0.5

= 0.4599

NORMAL DISTRIBUTION การแจกแจงปกติ



$$P\ (x < 215)$$

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 220}{7.5}$$

$$Z = \frac{215-220}{7.5} = -0.67$$

เปิดตาราง

P(Z<-0.67)= 0.2514

NORMAL DISTRIBUTION การแจกแจงปกติ

ถ้าคะแนนวิชาสถิติมีการแจกแจงปกติ มีค่าเฉลี่ย 74 และส่วนเบี่ยงเบนมาตรฐาน 7.9

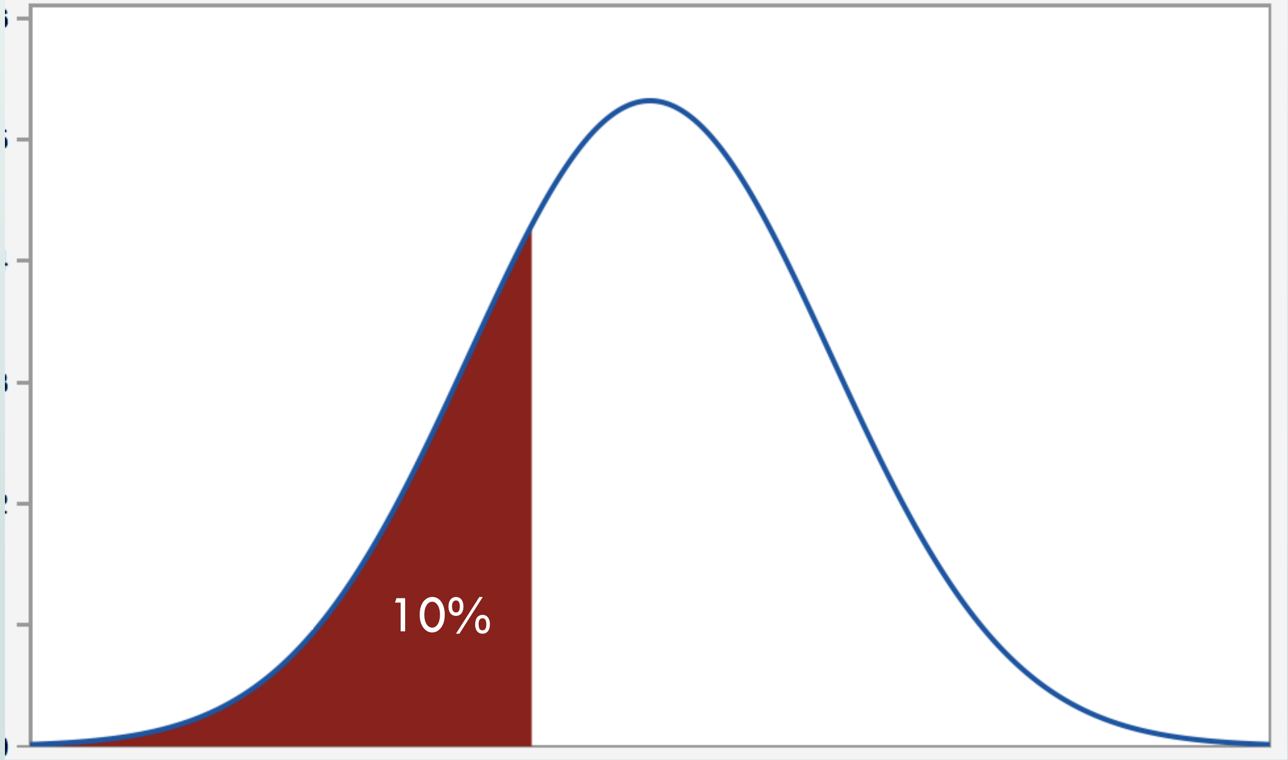
จงหา

ก. คะแนนต่ำสุดที่จะสอบผ่าน ถ้านักศึกษาสอบได้คะแนนต่ำที่สุดรวม 10% ได้เกรด F

ข. คะแนนสูงสุดสำหรับ เกรด B ถ้านักศึกษาสอบได้คะแนนสูงสุดรวม 5% ได้เกรด A

ตัวอย่าง

ข้อ ก $\mu = 74, \sigma = 9$



เกรด F Z

$P(Z < ???) = 0.1$

เปิดตารางที่ prob = 0.1
ได้ค่า Z ที่ -1.28

แทนค่า Z เพื่อหาค่า X

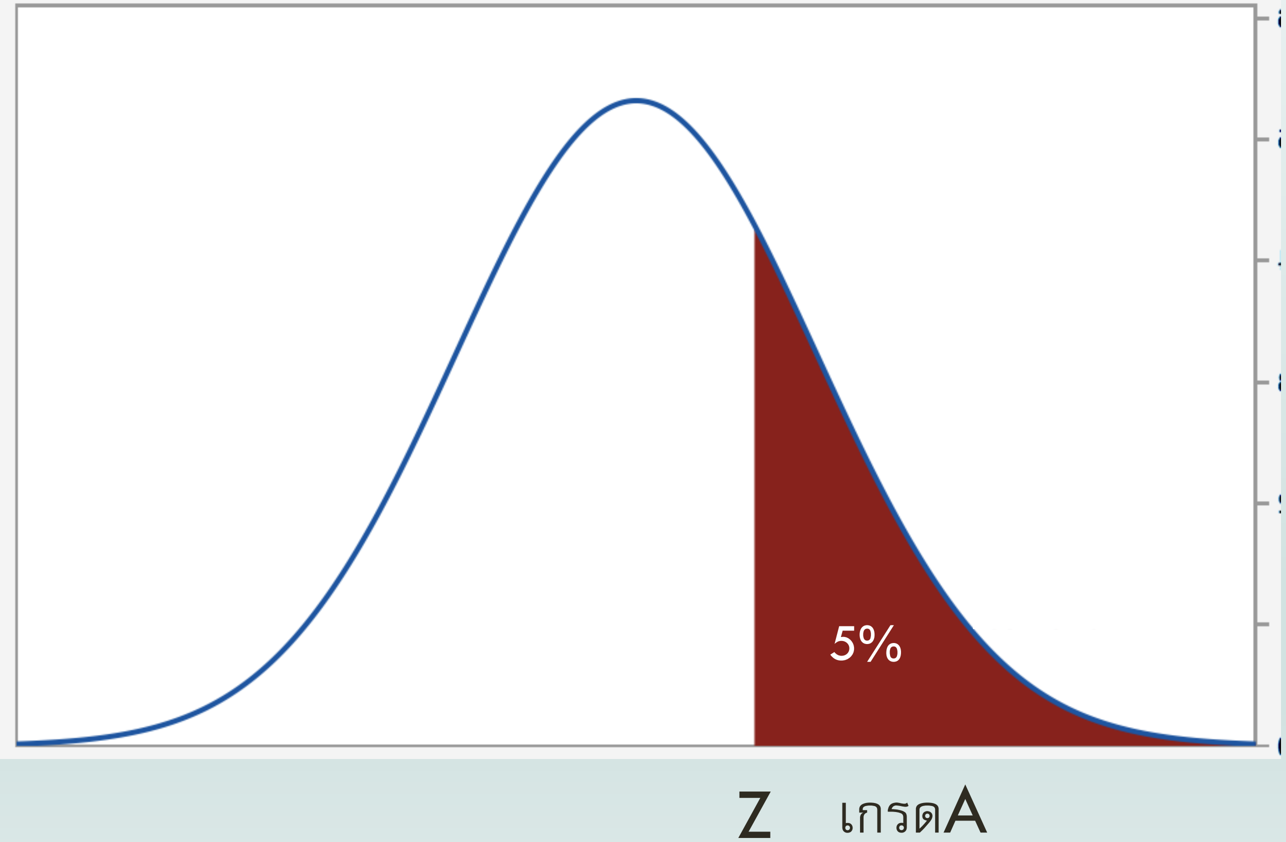
$$Z = \frac{X - \mu}{\sigma}$$

$$-1.28 = \frac{X - 74}{9}$$

X = 62.48

ตัวอย่าง

ข้อ ข $\mu = 74, \sigma = 9$



$P(Z < ???) = 1 - 0.05 = 0.95$

เปิดตารางที่ prob = 0.95

ได้ค่า Z ที่ 1.645

แทนค่า Z เพื่อหาค่า X

$$Z = \frac{X - \mu}{\sigma}$$

$$1.645 = \frac{X - 74}{9}$$

X = 88.805

NORMAL DISTRIBUTION การแจกแจงปกติ

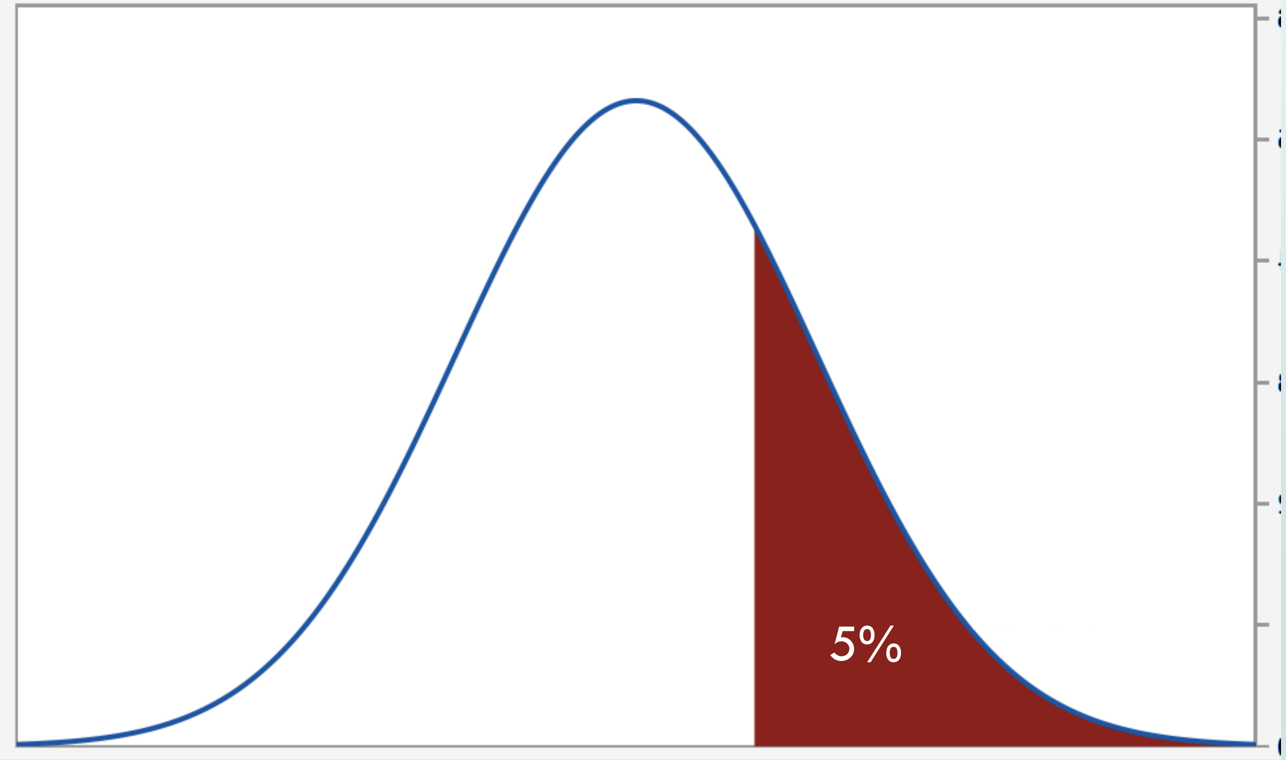
ความสูงของนักเรียนจำนวนหนึ่งมีการแจกแจงปกติ และมีความสูงเฉลี่ย 68.5 นิ้ว ถ้านักเรียนที่มีความสูงอย่างน้อย 71.2 นิ้ว มีอยู่ 12%

ก.จงคำนวณหาส่วนเบี่ยงเบนมาตรฐานของการแจกแจงนี้

ข. จงคำนวณหาความน่าจะเป็นที่นิสิตคนหนึ่งจะมีความสูงไม่เกิน 64 นิ้ว

ตัวอย่าง

ข้อ ก $\mu = 74\,,\sigma = 9$



Z เกรดA

$$P\,(x < 215)$$

$$Z = \frac{X-\mu}{\sigma} = \frac{X-220}{7.5}$$

$$Z = \frac{215-220}{7.5} = \text{-}0.67$$

เปิดตาราง

P(Z<-0.67)= 0.2514

NORMAL DISTRIBUTION การแจกแจงปกติ

ใช้ทำอะไร?

- อธิบายปรากฏการณ์ในชีวิตประจำวัน **e.g.**การแตกของเมล็ดป๊อบคอร์น , ส่วนสูงของประชากร , ผลสอบ, พฤติกรรมการจอดรถ

- ใช้ในสถิติเชิงอนุมาน

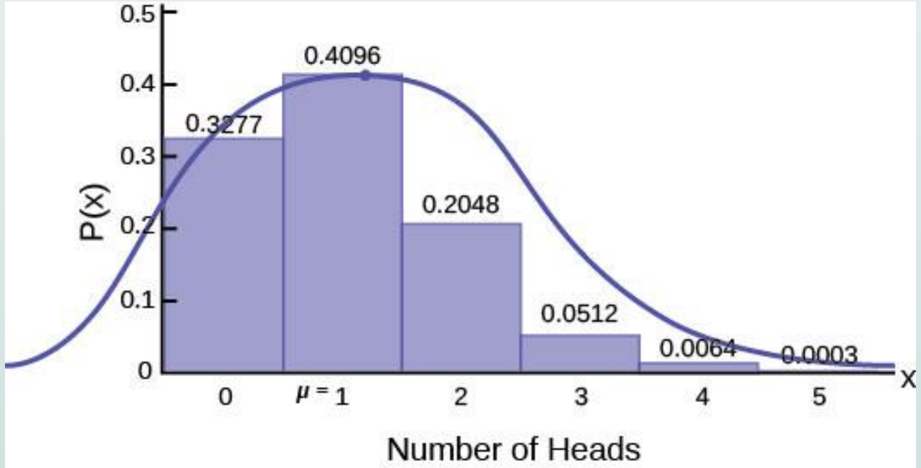
CENTRAL LIMIT THEOREM

The central limit theorem is a theorem in probability theory that establishes that the mean of a large number of independent and identically distributed random variables, when suitably rescaled, tends to a **normal distribution** [1] [2]. This holds true regardless of the shape of the original population distribution, provided the sample size is sufficiently large (usually n > 30) [3] [4] [5]. The central limit theorem also implies that the average of the sample means and standard deviations will equal the population mean and standard deviation [3] [4].

Pierre-Simon Laplace

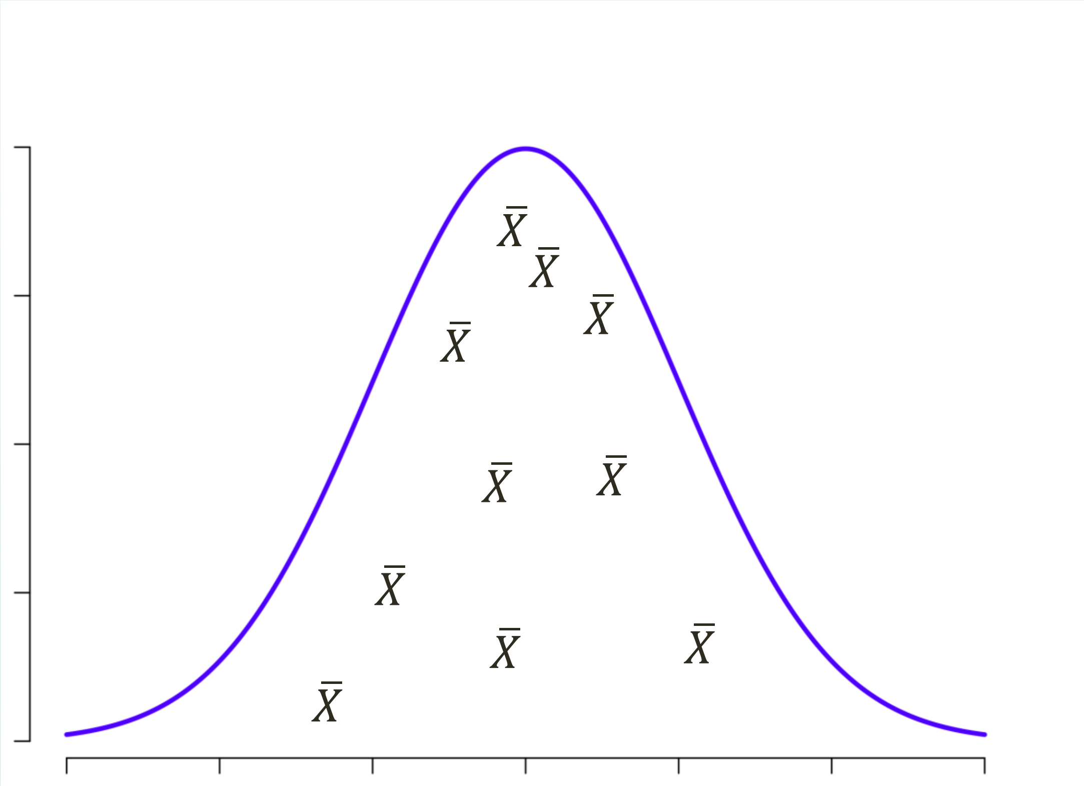
CENTRAL LIMIT THEOREM

Any Distribution



$$\bar{X} = \mu$$
$$S = \sigma$$

Normal Distribution



μ

CENTRAL LIMIT THEOREM

ค่าคลาดเคลื่อนมาตรฐาน Standard Error

$$\text{S.E.} = \frac{s}{\sqrt{n}}$$

n ยิ่งเยอะ ค่าคลาดเคลื่อนยิ่งน้อย

CENTRAL LIMIT THEOREM

เมื่อสุ่มตัวอย่าง **n** > 30

การแจกแจงความน่าจะเป็นของค่าเฉลี่ยจะเข้าใกล้การแจกแจงปกติ

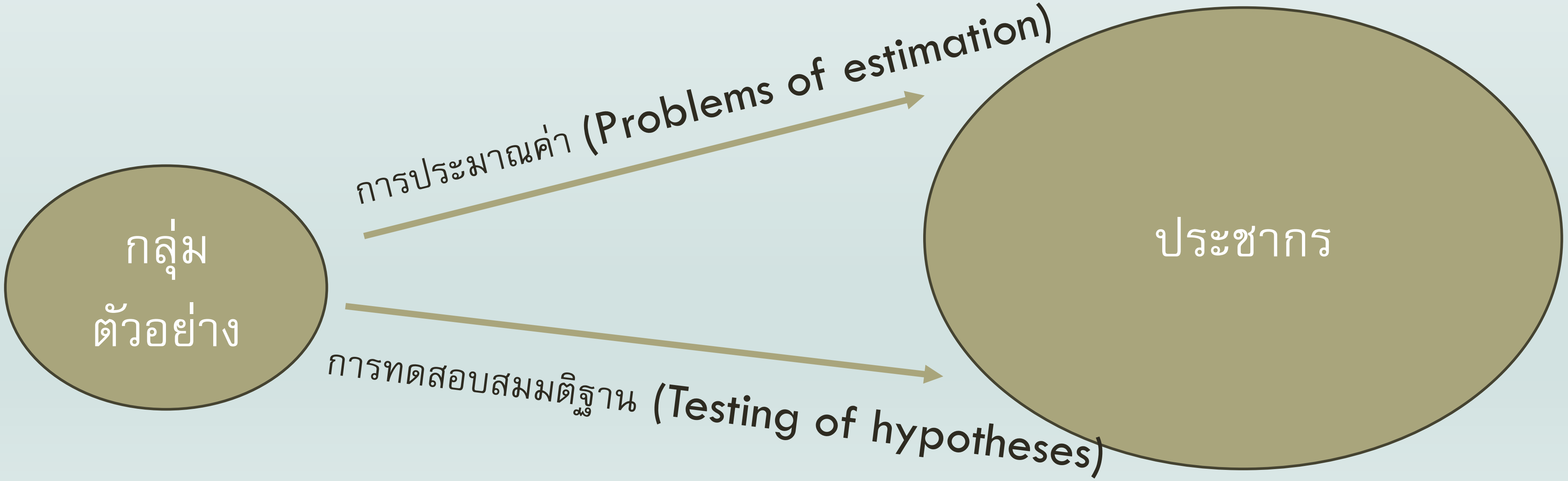
ถ้าเลือกกลุ่มตัวอย่างมาอย่างเหมาะสม → สามารถเป็นตัวแทนประชากรได้

STATISTICS

- สถิติเชิงพรรณนา (Descriptive Statistics)
  - ค่ากลาง
  - Measure of position
  - การกระจาย
- สถิติเชิงอนุมาน (Inferential Statistics)
  - การประมาณค่า
  - การทดสอบสมมติฐาน

การอนุมาน

สถิติเชิงอนุมาน (**Statistical inference**) คือ การใช้ข้อมูลที่สุ่มตัวอย่างมาจากประชากรไปประมาณสิ่งที่ต้องการศึกษามาจากประชากร



การอนุมาน

วัคซีนโควิด
ใช้ได้ดีหรือไม่

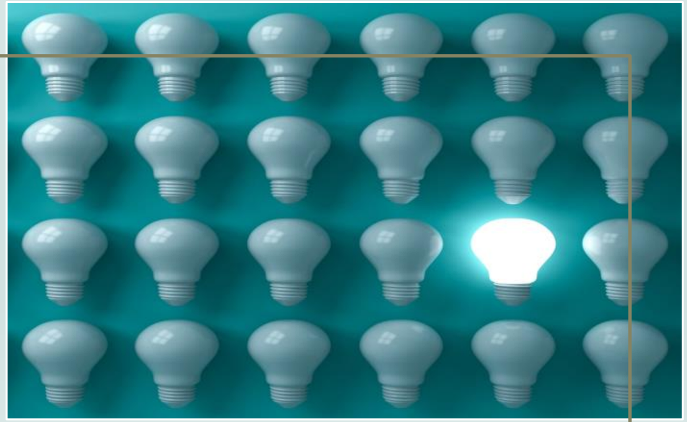
ลูกค้าสองกลุ่ม
มีพฤติกรรมเหมือนกันหรือไม่

แคมเปญโฆษณา
เพิ่มยอดขายได้จริงหรือไม่

การอนุมาน

ข้อมูล vs ความน่าจะเป็น

ข้อมูลคุยกับเรา



เข้าใจปรากฏการณ์ต่างๆ

การประมาณค่า

สัญลักษณ์ของค่าพารามิเตอร์และสถิติ

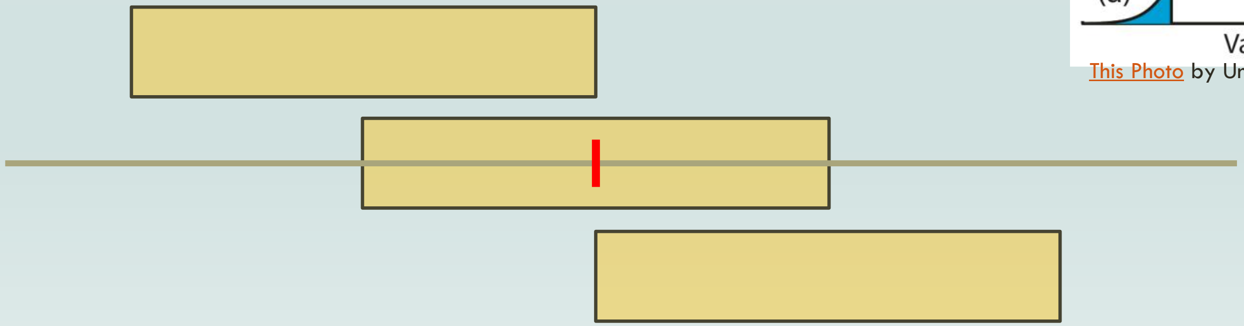
ประชากร	กลุ่มตัวอย่าง	ความหมาย
μ	$\bar{x}$	ค่าเฉลี่ย
σ	s	ส่วนเบี่ยงเบนมาตรฐาน
σ^2	s^2	ความแปรปรวน
p	$\hat{p}$	สัดส่วน
N	n	ขนาดของกลุ่มตัวอย่าง

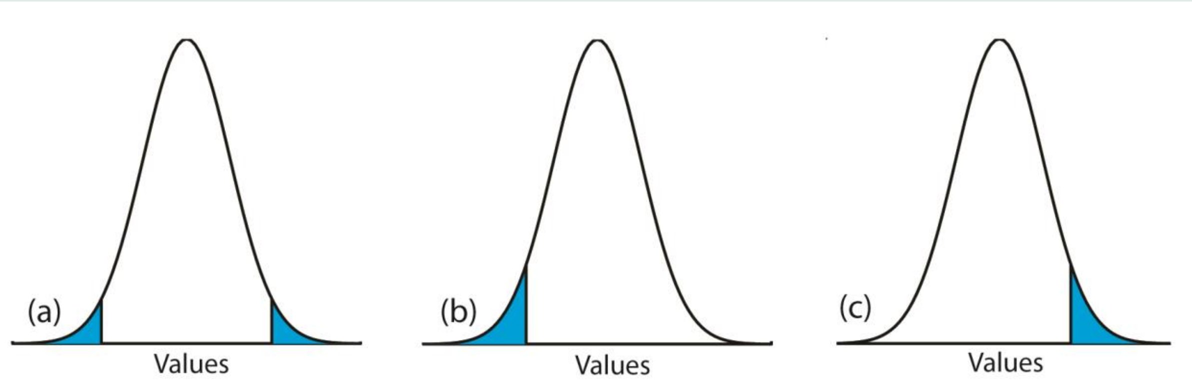
การประมาณค่า

การประมาณค่าแบบจุด

การประมาณค่าแบบช่วง

- Two tailed interval estimate
- One-tailed interval estimate





การประมาณค่า

ร้านชานม Bearhouse ต้องการทราบอายุลูกค้าที่เข้ามาซื้อชานม จึงทำการสุ่มตัวอย่างมา 50 คน (n=50)

หาค่าเฉลี่ยของอายุได้เท่ากับ 24.5 ปี ($\bar{x}$ =24.5)

ค่าประมาณแบบจุด = 24.5 ปี

ค่าประมาณแบบช่วง = 18.5-30.5 ปี

CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

ค่าประมาณแบบช่วง เรียกอีกชื่อหนึ่งว่า

ช่วงความเชื่อมั่น

(Confidence interval)

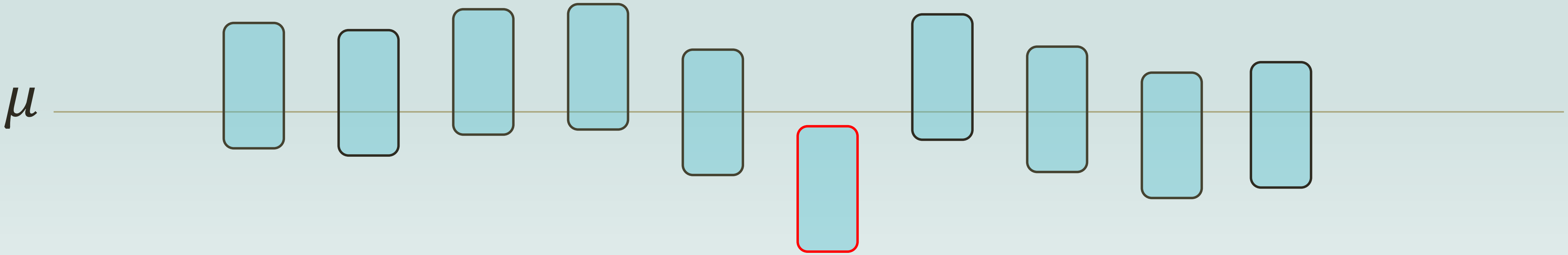
ความกว้างของช่วงความเชื่อมั่นที่ประมาณได้ ขึ้นกับ

- การกระจายของค่าที่จะทราบ (s,σ)
- ขนาดของตัวอย่างที่สุ่มมา
- ระดับความเชื่อมั่น $(1-\alpha) * 100\%$

CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

การประมาณค่าแบบช่วงของค่า μ ที่ระดับความเชื่อมั่น 90%

ในการประมาณค่าพารามิเตอร์ซ้ำๆกัน 100 ครั้ง
จะมี 90 ครั้งที่ช่วงความเชื่อมั่นที่สร้างขึ้น ครอบคลุมค่าพารามิเตอร์
และมี 10 ครั้งที่ไม่ครอบคลุมค่าพารามิเตอร์



CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

ค่าโอกาสที่จะยอมให้ผิดพลาดได้ เรียกว่า

ระดับนัยสำคัญ (Significant level)

สัญลักษณ์ α

CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

ระดับความเชื่อมั่น (1-α)x100%

- โอกาสที่ค่าพารามิเตอร์ที่ประมาณจะ อยู่ ในช่วงที่เราสร้างขึ้น
- อยู่ในรูป % สูงๆ เช่น 90%, 95%, 99%

ระดับนัยสำคัญ (α)

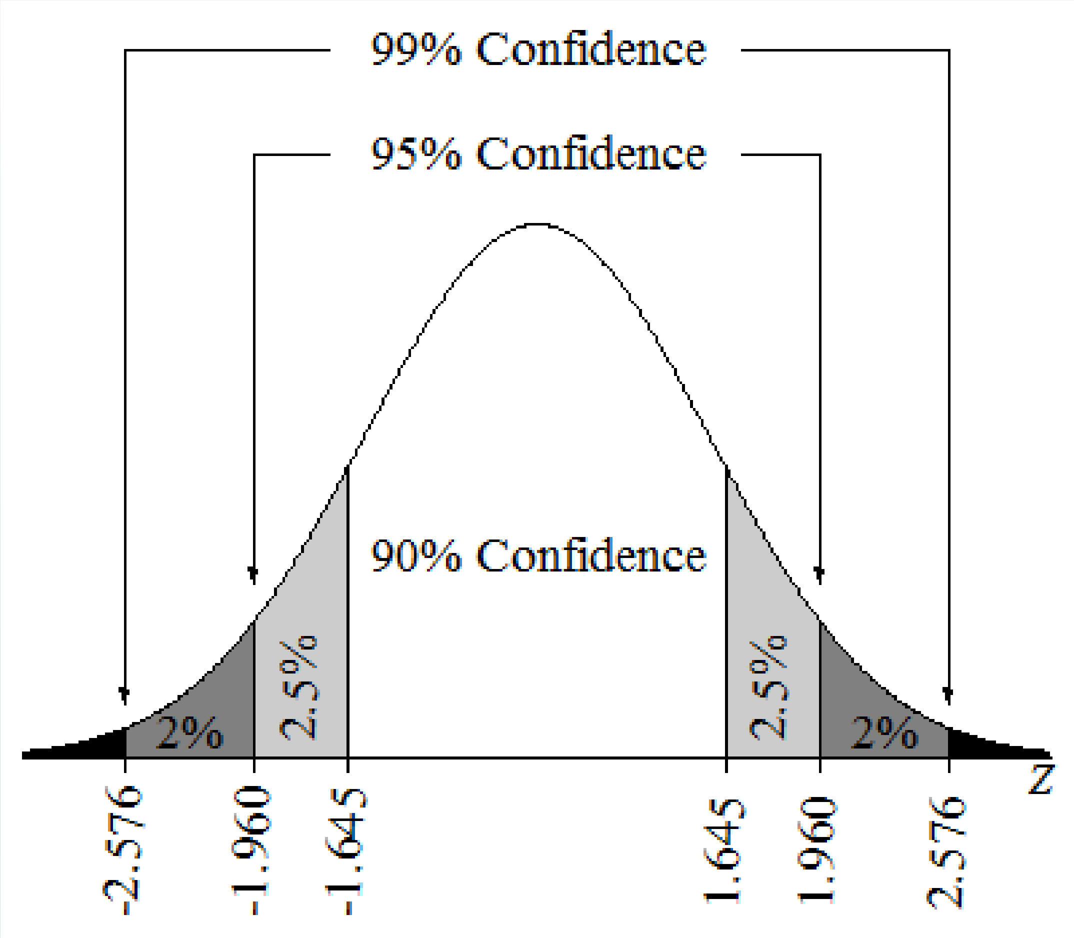
- โอกาสที่ค่าพารามิเตอร์ที่ประมาณจะ ไม่อยู่ ในช่วงที่เราสร้างขึ้น
- อยู่ในรูปทศนิยม เช่น 0.05 , 0.01, 0.02
- สามารถคำนวณได้จาก ระดับความเชื่อมั่น

ระดับนัยสำคัญ = (100-ระดับความเชื่อมั่น)/100

CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

ระดับความเชื่อมั่น	ระดับนัยสำคัญ (α)
99%	
95%	
92%	
90%	
88%	

CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

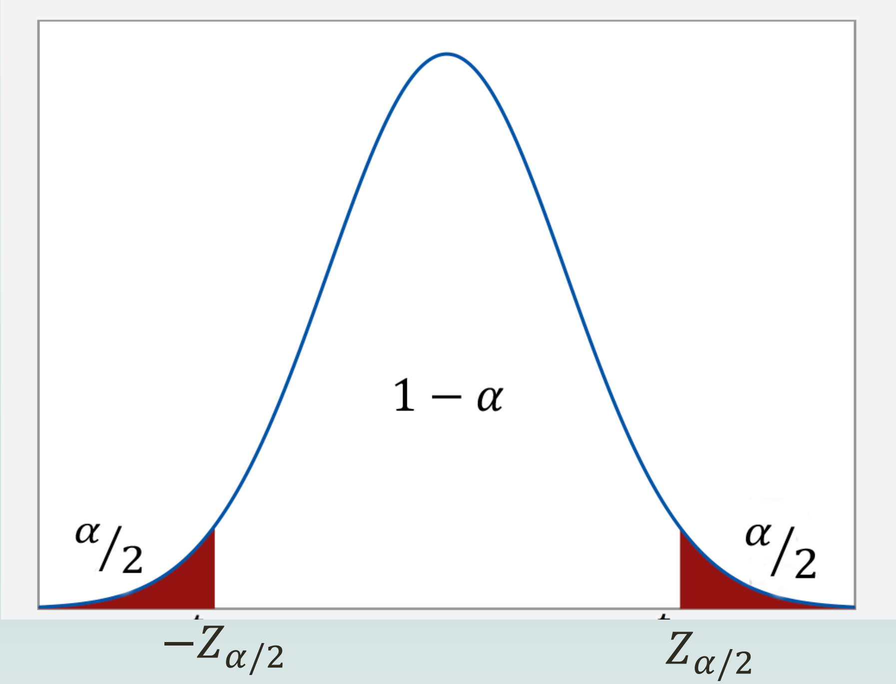


ระดับความเชื่อมั่น 95% หรือระดับนัยสำคัญ 0.05

พื้นที่อยู่นอกช่วงประมาณค่ารวมกันเท่ากับ 0.05 (α)
แบ่งซ้ายขวา ฝั่งละ 0.025 ($\alpha/2$)

CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

ค่า $Z_{\alpha/2}$ คือ ค่า Z ณ จุดที่ทำให้พื้นที่ฝั่งขวามือของโค้งปกติมีค่าเท่ากับ $\alpha/2$



CONFIDENCE INTERVAL ช่วงความเชื่อมั่น

ระดับความเชื่อมั่น	α	$\alpha/2$	$z_{\alpha/2}$	z_{α}
90%				
92%				
94%				
95%				
96%				
98%				
99%				

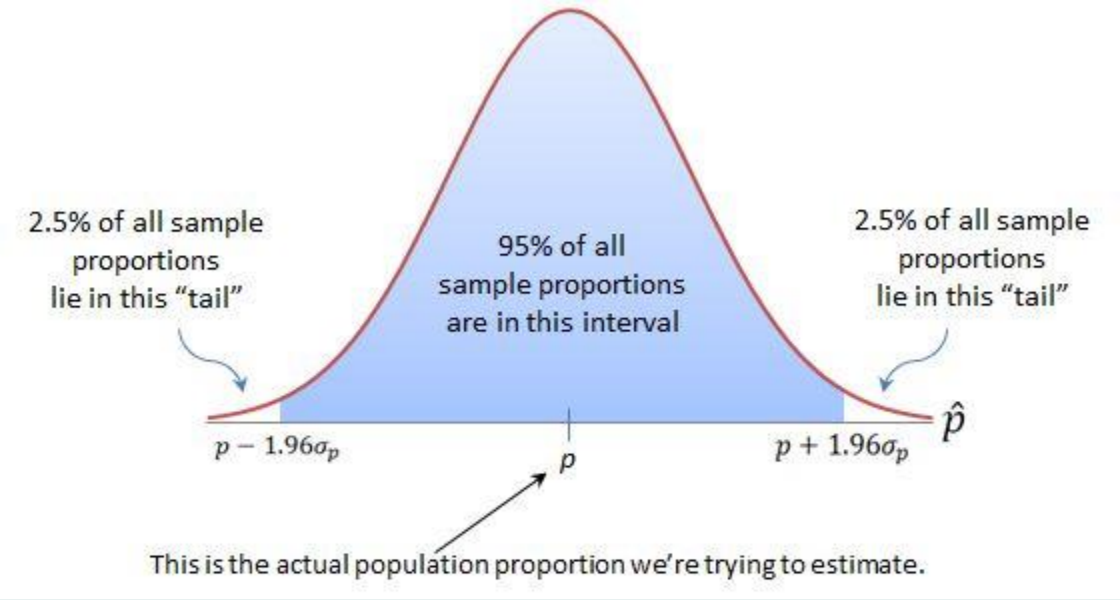
การประมาณช่วงความเชื่อมั่น

การประมาณค่า
➢ ค่าเฉลี่ยประชากรหนึ่งกลุ่ม (μ) ➢ ผลต่างของค่าเฉลี่ยประชากรสองกลุ่ม ($\mu_1 - \mu_2$) ➢ สัดส่วน (p)
➢ ผลต่างของสัดส่วน ($p_1 - p_2$) ➢ ค่าความแปรปรวน (σ^2) ➢ อัตราส่วนของค่าความแปรปรวน (σ_1^2/σ_2^2)

การประมาณช่วงความเชื่อมั่น

ความกว้างของช่วงที่ประมาณได้ ขึ้นอยู่กับ

- การกระจายของค่าที่ต้องการทราบ
- ขนาดตัวอย่าง
- ระดับความเชื่อมั่น



ช่วงความเชื่อมั่นของค่าเฉลี่ย (μ)

เมื่อทราบความแปรปรวนของประชากร

$$\mu = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

เมื่อไม่ทราบความแปรปรวนของประชากร

N > 30

$$\mu = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

N < 30

$$\mu = \bar{x} \pm t_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

ช่วงความเชื่อมั่นของค่าเฉลี่ย (μ)

ตัวอย่าง **:**

สุ่มตัวอย่างเครื่องดื่มชนิดหนึ่งจำนวน 36 ราย พบว่า มีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของจำนวนน้ำตาล เป็น 22 g และ 4 g ตามลำดับ จงหาขอบเขตที่เชื่อมั่นได้ 90% ของปริมาณน้ำตาลที่แท้จริงที่มีอยู่ในเครื่องดื่มชนิดนี้

ช่วงความเชื่อมั่นของค่าเฉลี่ย (μ)

n=36

$\bar{x} = 22$ g , s= 4 g

$\alpha = 0.10 , \frac{\alpha}{2} = 0.05$

$z_{\alpha/2} = 1.645$

$$\mu = \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$= 22 \pm (1.645) \frac{4}{\sqrt{36}}$$

$$= 22 \pm 1.0967$$

$$22 - 1.0967 \leq \mu \leq 22 + 1.0967$$

$$20.9033 \leq \mu \leq 23.0967$$

ช่วงความเชื่อมั่นของผลต่างของค่าเฉลี่ย$(\mu_1 - \mu_2)$

เมื่อทราบความแปรปรวนของประชากรของทั้งสองกลุ่ม σ_1^2, σ_2^2

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

เมื่อไม่ทราบความแปรปรวนของประชากร

N > 30

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

ช่วงความเชื่อมั่นของผลต่างของค่าเฉลี่ย$(\mu_1 - \mu_2)$

ตัวอย่าง :

สุ่มตัวอย่างความพึงพอใจของลูกค้าต่อผลิตภัณฑ์ A และ B โดยกลุ่มลูกค้า 50 คนทั้งสองผลิตภัณฑ์ ผลการทดสอบคือ

ผลิตภัณฑ์ A ได้คะแนนเฉลี่ย 76 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 6

ผลิตภัณฑ์ B ได้คะแนนเฉลี่ย 82 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 8

จงหาช่วงความเชื่อมั่น 96% สำหรับผลต่างคะแนนความพึงพอใจของทั้งสองผลิตภัณฑ์

ช่วงความเชื่อมั่นของผลต่างของค่าเฉลี่ย$(\mu_1 - \mu_2)$

n_1=50 , n_2=50

$\bar{x}_1 = 82 , \bar{x}_2 = 76$

s_1=8 , s_2=6

$\alpha = 0.04 , \frac{\alpha}{2} = 0.02$

$z_{\alpha/2}$ = 2.05

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= (82 - 76) \pm (2.05)\sqrt{\frac{8^2}{50} + \frac{6^2}{50}}$$

$$= \mathbf{6 \pm 2.8991}$$

$$6 - 2.8991 \leq \mu_1 - \mu_2 \leq 6 + 2.8991$$

$$3.1009 \leq \mu_1 - \mu_2 \leq 8.8991$$

ช่วงความเชื่อมั่นของสัดส่วน (P)

ค่าสัดส่วน คำนวณจาก ความถี่ของสิ่งที่เราสนใจ หารด้วยความถี่ทั้งหมด

เก็บข้อมูล 500 คน
(n=500)

ขับรถมาเรียน
320 คน
(x=320)

สัดส่วนผู้ที่ขับรถมาเรียน
320/500 = 0.64
(p=0.64)

ไม่ขับรถมาเรียน
180 คน
(x=180)

สัดส่วนผู้ที่ไม่ขับรถมาเรียน
180/500 = 0.36
q = 0.36 = 1-p

ช่วงความเชื่อมั่นของสัดส่วน (P)

$$p = \hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$\hat{p}$ คือสัดส่วนของสิ่งที่เราสนใจ

$\hat{q}$ คือสัดส่วนของสิ่งที่เราไม่สนใจ (1- $\hat{p}$)

ช่วงความเชื่อมั่นของสัดส่วน (P)

ตัวอย่าง

สุ่มผู้ขับรถแถวถนนฉลองกรุงจำนวน 200 คน พบว่า 175 คน ใช้หูฟังในการโทรศัพท์ จงประมาณค่าสัดส่วนผู้ขับรถแถวถนนฉลองกรุงที่จะไม่ใช้หูฟังในการโทรศัพท์ ที่ระดับความเชื่อมั่น 98%

ช่วงความเชื่อมั่นของสัดส่วน (P)

$\hat{p} = 25/200 = 0.125$

$\hat{q} = 1 - \hat{p} = 1 - 0.125 = 0.875$

$\alpha = 0.02 , \frac{\alpha}{2} = 0.01$

$z_{\alpha/2} = 2.33$

$$\mathrm{p} = \hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= 0.125 \pm (2.33) \sqrt{\frac{0.125x0.975}{200}}$$

$$= 0.125 \pm 0.0545$$

$$0.125 - 0.0545 \leq \hat{p} \leq 0.125 + 0.0545$$

$$0.0705 \leq \hat{p} \leq 0.1795$$

$$7.05\% \leq \hat{p} \leq 17.95\%$$

WRAP UP

- Joint distributions, covariance
- Normal Distribution
- Central limit theorem
- การประมาณค่า

Q & A