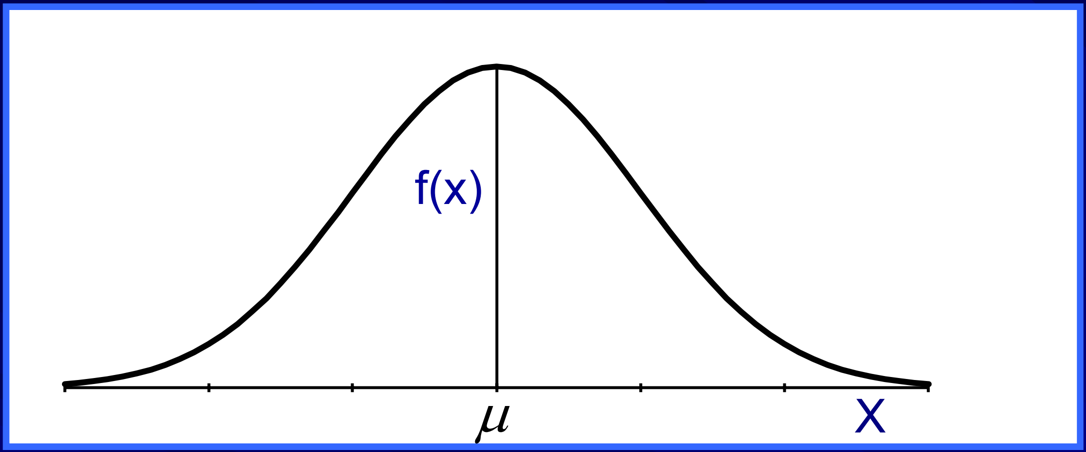
The Normal Distribution

Summer 2003

Normal Distribution: Characteristics



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Where:

μ = mean of X

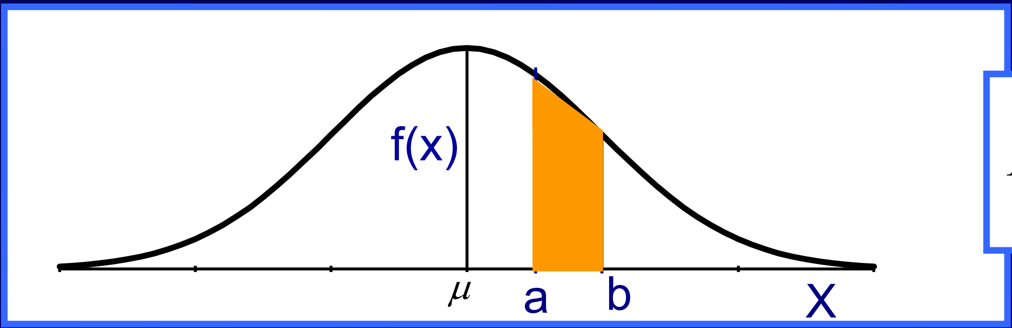
σ = standard deviation of X

π = 3.14159 . . .

e = 2.71828 . . .

- "Famous" "bell shaped" PDF, unimodal (only one hump).
- Area under the curve sums to 1.
- Symmetrical Distribution: Area to right/left of mean is 1/2.
- Asymptotic to the Horizontal Axis.
- A Family of Curves: The Normal R.V. X is given by two parameters μ and σ. [r.v. X N(μ,σ)].

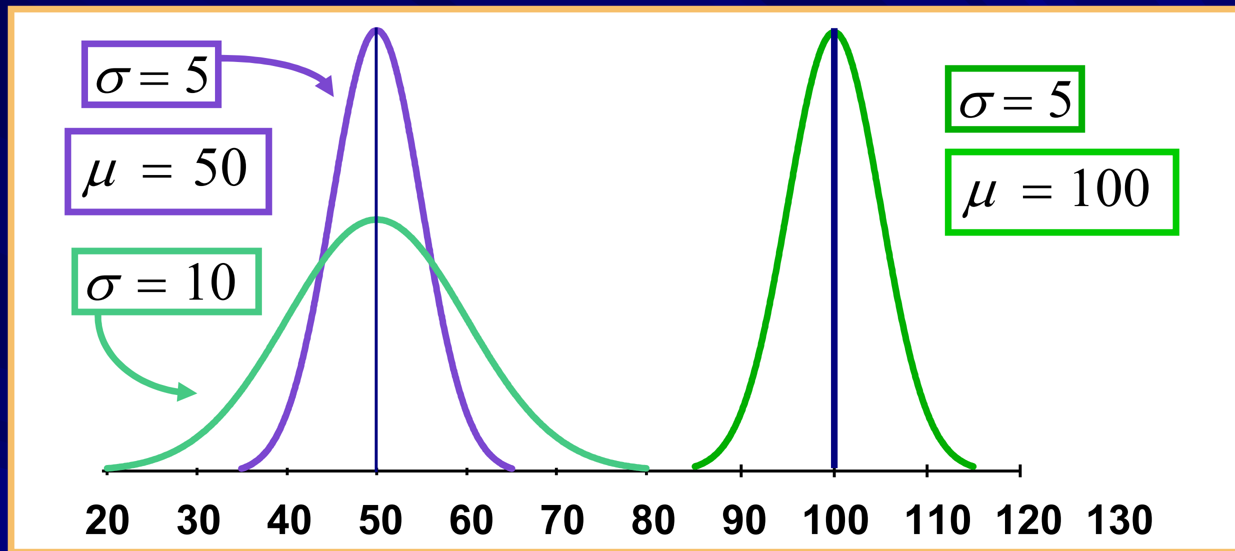
Normal Distribution: calculating probabilities



$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

- The integral of f(x) for the normal distribution does not have a closed form, i.e. it can't be solved.
- Thus, the area under the normal curve must be calculated using numerical methods.
- We will use the table in page 518.
- Table gives the CDF for the standard normal r.v. Z N(0,1)

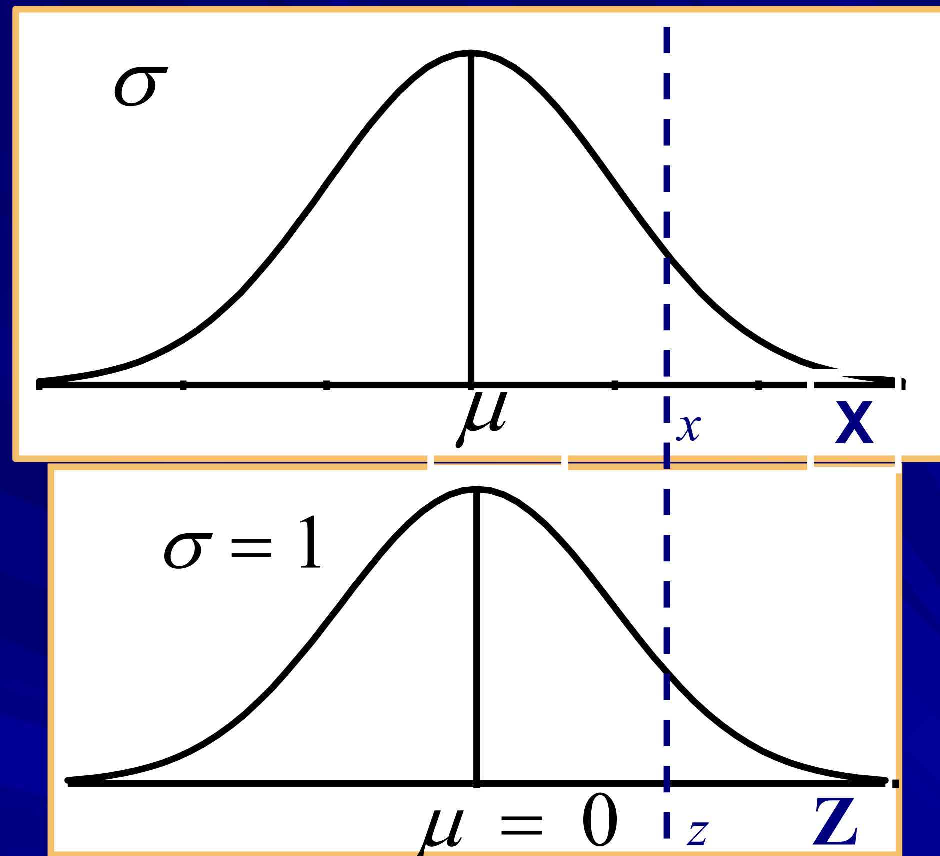
Normal Curves for Different Means and Standard Deviations



- There is an infinite number of normal curves.
- How can we calculate normal probabilities using just the CDF table for r.v. Z N(μ=0,σ=1)?
- We will use a transformation that converts value x from r.v. X N(μ,σ) into value z from r.v. Z N(μ=0,σ=1)

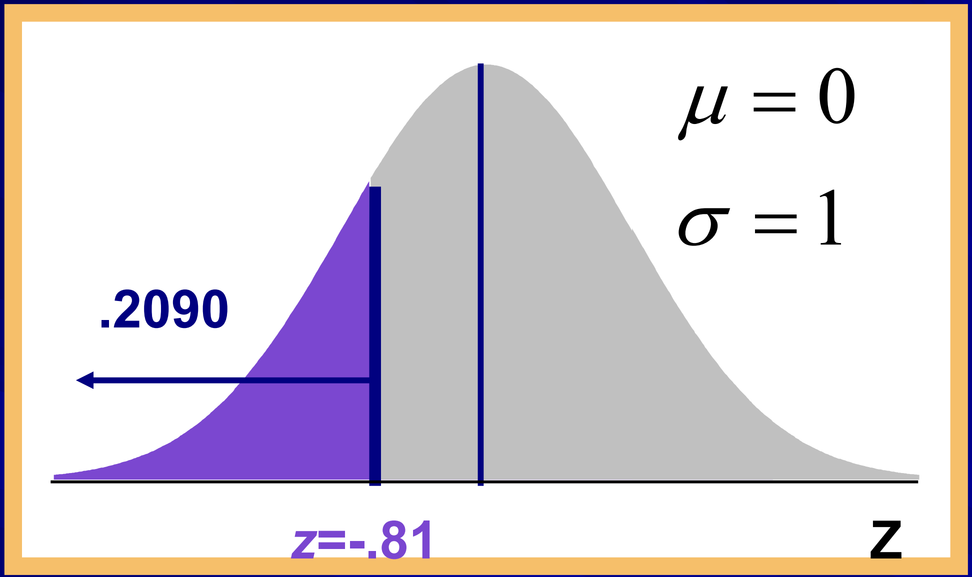
Standardized Normal Distribution

- Value x from RV X N(μ,σ):
- Value z from RV Z N(0,1):
- z Score transformation: computed by the Z Formula. z represents the number of standard deviations an x value is away from the mean.



$$z = \frac{x - \mu}{\sigma}$$

Using the CDF table for Z N(0,1)



For z=-0.81, $P(Z \leq -0.81) = F(z) = 0.2090$

(see table next two slides)

Cumulative Distribution Function of the Standard Normal Distribution

Example: If Z is standard Normal random variable, then F(1.00) = P(Z ≤ 1.00) = .8413

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-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

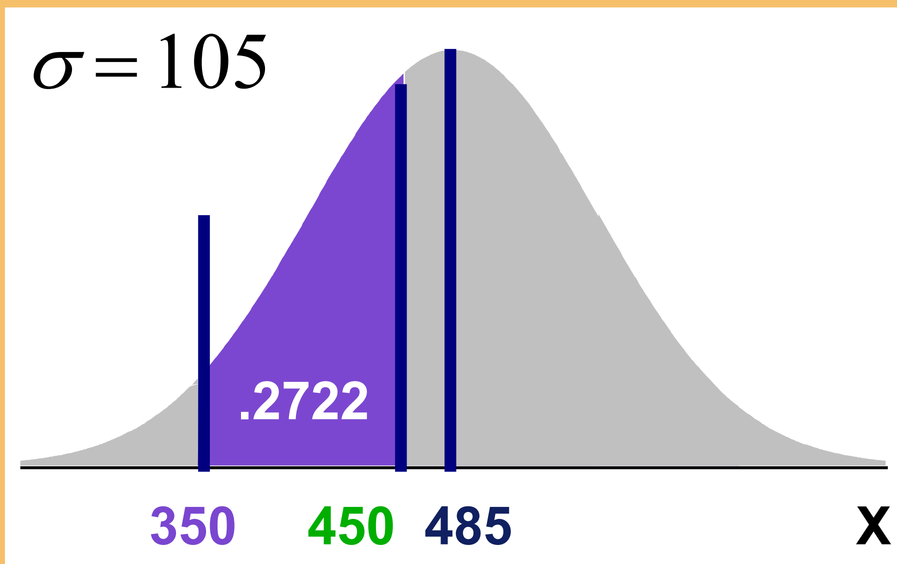
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	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

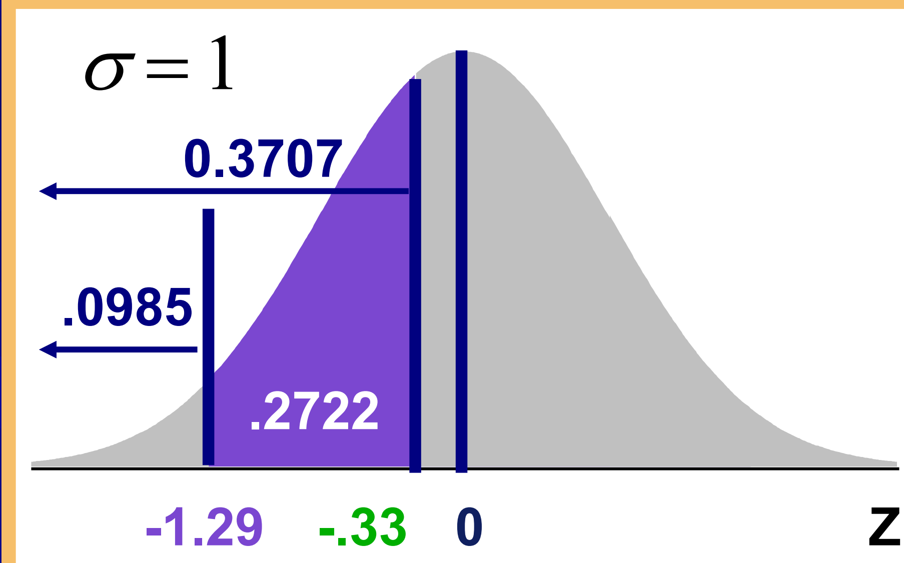
Example

X is normally distributed with $\mu = 485$, and $\sigma = 105$

$$P(350 \le x \le 450) = P(-1.29 \le z \le -0.33)$$

$$= P(z \le -0.33) - P(z \le -1.29) = 0.3707 - 0.0985 = 0.2722$$

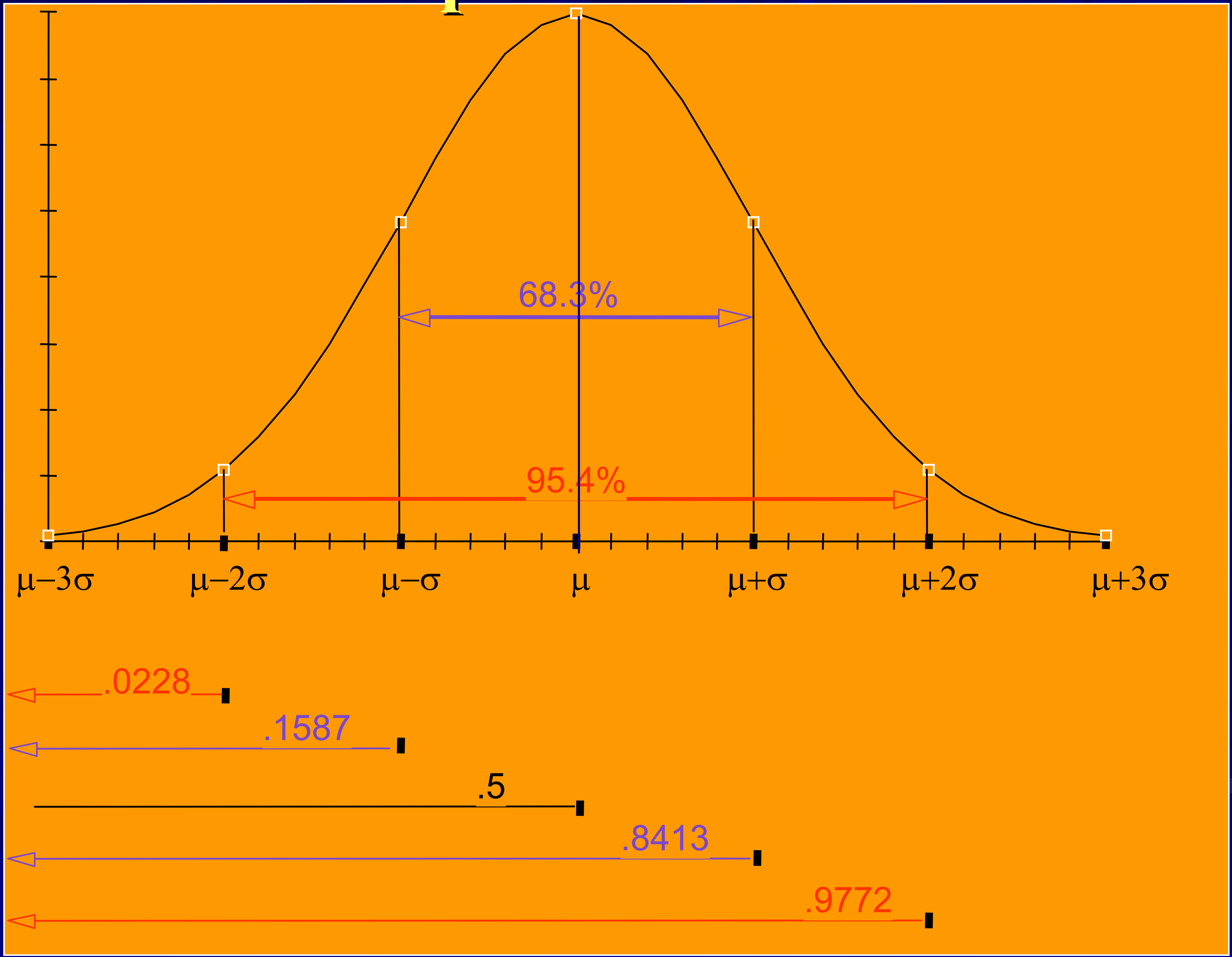




$$Z = \frac{X - \mu}{\sigma} = \frac{350 - 485}{105} = -1.29$$

$$Z = \frac{X - \mu}{\sigma} = \frac{450 - 485}{105} = -0.33$$

Some helpful rules of thumb



A randomly selected value from a Normal distribution will fall within:

- 1 standard deviation 68.3% of the times…

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.683$$

- 2 standard deviations 95.4% of the times…

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.954$$

- 3 standard deviations 99.7% of the times

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.997$$

essentially, it falls within three standard deviations most of the time!

Six Sigma

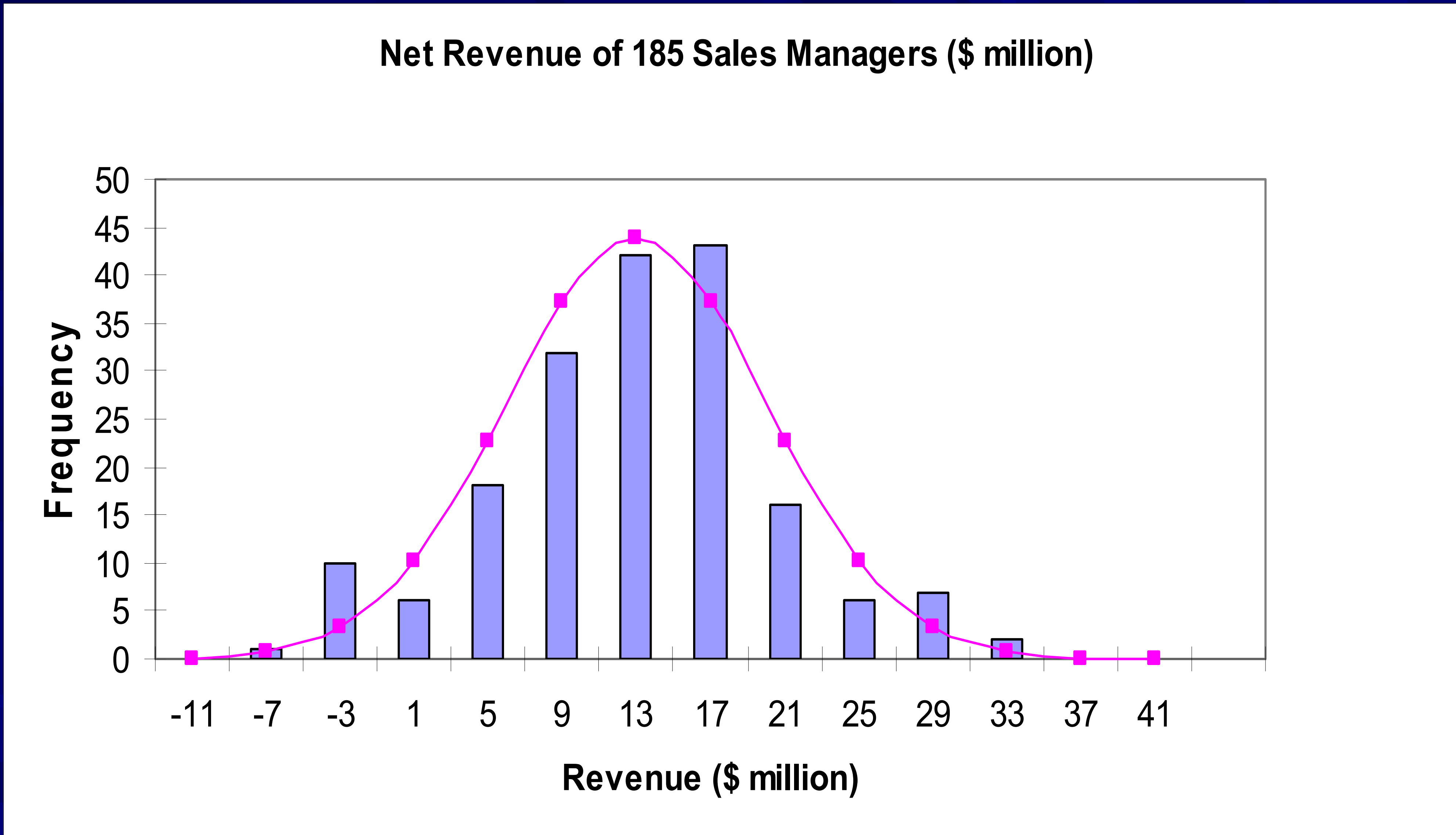
- The standard table of CDF for Z on p. 518 gives us values for Z between –3 and 3.
- How many of you are familiar with "Six Sigma" concepts in TQM?
- "Three Sigma" quality means that 1 unit of every 300 is expected to fall outside the tolerance limit; this may not be good enough!
- Six Sigma means that 1 unit of every 300,000 falls outside the tolerance limit; this is a very difficult goal that some companies achieve routinely

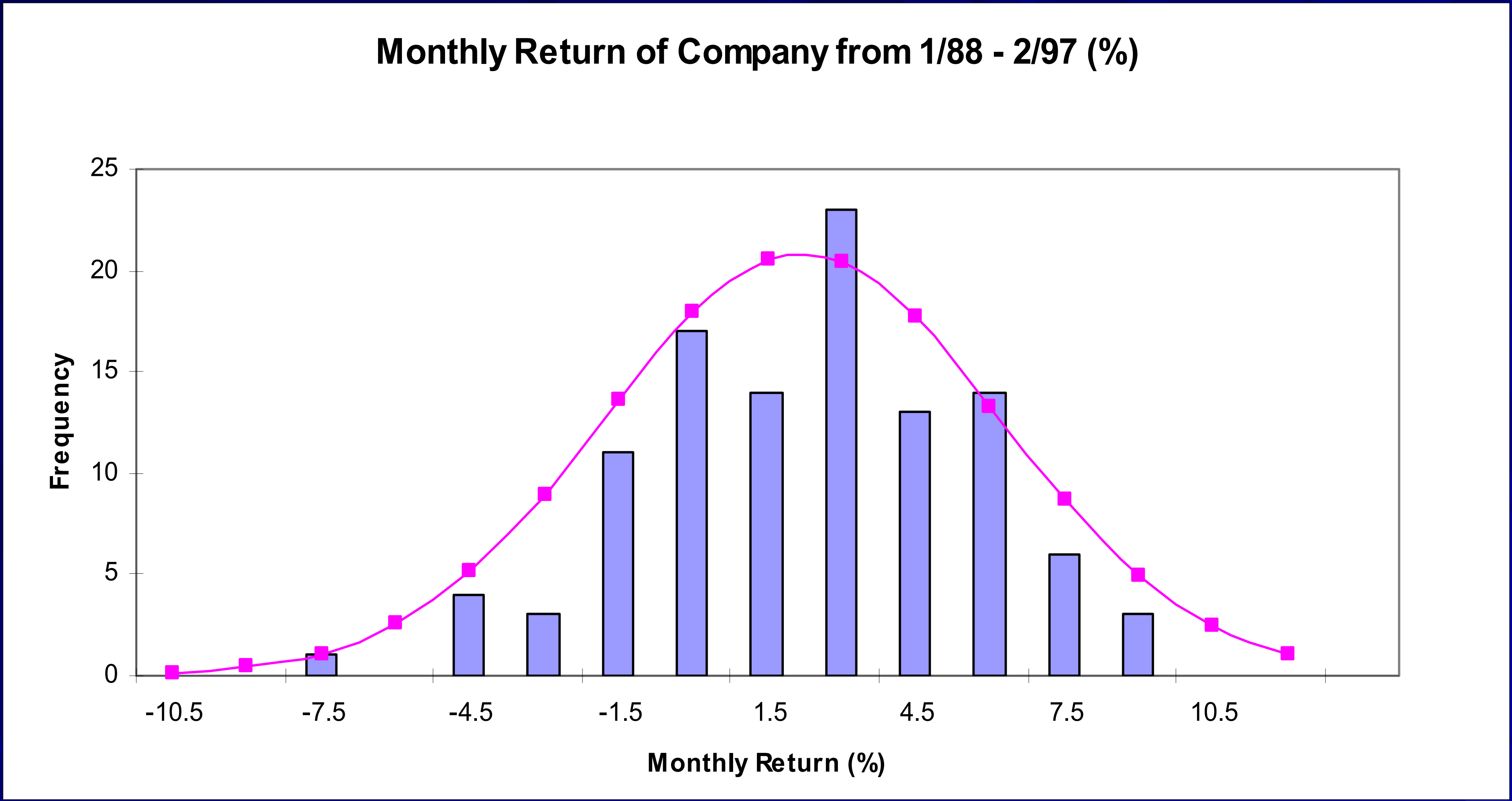
Many phenomena obey the Normal distribution:

- GMAT scores of Sloan class of 2001
- Height of a group of people (e.g., Sloan faculty)
- Stock return over short periods of time
- Width of steel plate from a production process

But ... Some phenomena are not Normally distributed:

- Stock returns over longer periods of time are not Normal
- Income distributions are typically not Normal


Net Revenue of 185 Sales Managers ($ million)
Frequency
50
45
40
35
30
25
20
15
10
5
0
-11
-7
-3
1
5
9
13
17
21
25
29
33
37
41
Revenue ($ million)


Monthly Return of Company from 1/88 - 2/97 (%)
Frequency
25
20
15
10
5
0
-10.5
-7.5
-4.5
-1.5
1.5
4.5
7.5
10.5
Monthly Return (%)

Monthly Rate of Return of a Company (%)

(1/88 - 2/97)

Month	Return	Month	Return	Month	Return	Month	Return	Month	Return
Jan-88	3.58147	Nov-89	-1.129413	Sep-91	0.653083	Jul-93	-0.444741	May-95	-0.078501
Feb-88	2.41552	Dec-89	-1.63933	Oct-91	4.804913	Aug-93	-1.560677	Jun-95	-1.397605
Mar-88	2.472095	Jan-90	2.16927	Nov-91	0.461484	Sep-93	3.691086	Jul-95	5.691725
Apr-88	1.999385	Feb-90	1.733777	Dec-91	2.037275	Oct-93	2.281874	Aug-95	3.821112
May-88	-1.406636	Mar-90	6.403131	Jan-92	-0.071025	Nov-93	-3.484922	Sep-95	2.898039
Jun-88	7.500428	Apr-90	-5.59515	Feb-92	5.113477	Dec-93	3.683833	Oct-95	-2.645249
Jul-88	4.609917	May-90	-4.775914	Mar-92	-7.681604	Jan-94	-1.785277	Nov-95	-1.062747
Aug-88	4.184928	Jun-90	6.436919	Apr-92	4.162241	Feb-94	3.663833	Dec-95	3.101275
Sep-88	-2.135523	Jul-90	-4.071401	May-92	1.897831	Mar-94	1.120663	Jan-96	1.468295
Oct-88	4.844958	Aug-90	1.303258	Jun-92	2.122022	Apr-94	0.036396	Feb-96	0.728303
Nov-88	5.001955	Sep-90	0.245184	Jul-92	-3.516173	May-94	2.995746	Mar-96	7.248898
Dec-88	3.539133	Oct-90	-1.469709	Aug-92	2.078211	Jun-94	4.64556	Apr-96	-1.657884
Jan-89	3.228681	Nov-90	1.858332	Sep-92	5.263744	Jul-94	-2.690679	May-96	2.663203
Feb-89	5.323143	Dec-90	2.773697	Oct-92	0.860343	Aug-94	-0.750704	Jun-96	7.87292
Mar-89	-2.186619	Jan-91	5.720819	Nov-92	2.300328	Sep-94	-4.831575	Jul-96	7.00066
Apr-89	-1.478049	Feb-91	5.941408	Dec-92	-2.10963	Oct-94	0.235075	Aug-96	5.933959
May-89	-5.699251	Mar-91	-1.119768	Jan-93	-2.380887	Nov-94	5.063136	Sep-96	0.897587
Jun-89	6.651669	Apr-91	2.764807	Feb-93	1.517817	Dec-94	-0.909588	Oct-96	6.528705
Jul-89	8.76046	May-91	2.289069	Mar-93	1.546748	Jan-95	1.74575	Nov-96	0.768207
Aug-89	-0.249731	Jun-91	-2.196387	Apr-93	2.023996	Feb-95	3.992006	Dec-96	-1.436122
Sep-89	3.002285	Jul-91	3.331804	May-93	3.353796	Mar-95	2.742802	Jan-97	0.439757
Oct-89	-1.196443	Aug-91	-0.462699	Jun-93	-0.729763	Apr-95	4.562742	Feb-97	0.565258

Using Excel To Calculate Normal Probabilities

Use the function wizard icon or the Insert_Function Command to choose the Statistical function NORMDIST to calculate the cumulative probability for a normal random variable with given mean and sd.

Use NORSDIST to calculate cumulative probability for the standard normal random variable with mean=0 and sd=1.

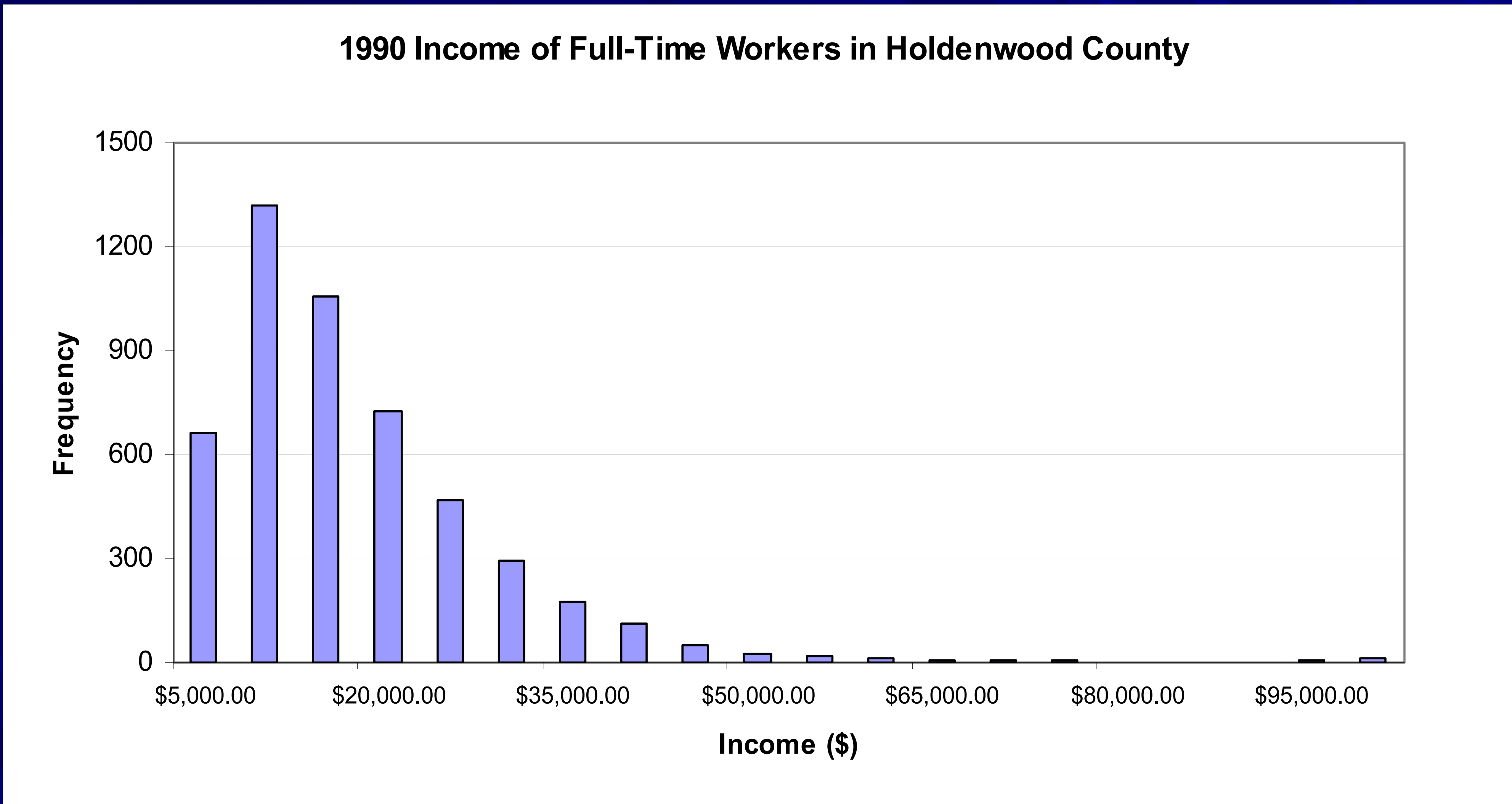
Over long periods of time stock returns follow a lognormal distribution. Why?

- If total stock returns over a number of years were calculated arithmetically ($r_1 + r_2 + \ldots$); then, by the CLT total stock return over long periods of time would follow a normal distribution.
- But Stocks returns are calculated geometrically: Total accumulated return: $R = (1+r_1)(1+ r_2)\ldots$

  Taking the log of both sides: $\log(R) = \log(1+r_1)+ \log(1+ r_2)+$

  By the CLT again, the log of the accumulated return follows a normal distribution. Thus, the accumulated return follows a lognormal distribution.

Consider the distribution of income



Let's Try A Practice Problem

Two retail chains in Boston/Cambridge (Circuit City and CompUSA) are planning to sell the Apple iMac computer. Demands at the two retail chains are random variables, and we are given their means, standard deviations, and the correlation between them. Suppose Circuit City and CompUSA are considering a merger. What is the distribution of the demand for the iMac computer from the combined company? What is the 98th percentile of the demand for iMacs at the merged company?

Structuring the problem: What are we trying to find out? What kind of problem is this? What do we know that would be useful?

Two parts to the problem. The goal is to find out something about the combined demand distribution of the merged company. But we know about the separate demands, so the first part is to create a distribution of the sum of two random variables, and the second part is to read information from that distribution.

Let X = daily demand at Circuit City, and Y = daily demand at CompUSA. Suppose we are told in the problem that X and Y are Normally distributed with:

μ_X	=	800	μ_Y	=	160
σ_X	=	500	σ_Y	=	100

CORR(X,Y) = 0.23

Answer…

If we let W = X + Y, then W is *normally distributed* with parameters:

$\mu_w = \mu_X + \mu_Y$ = 800+160 =960 computers

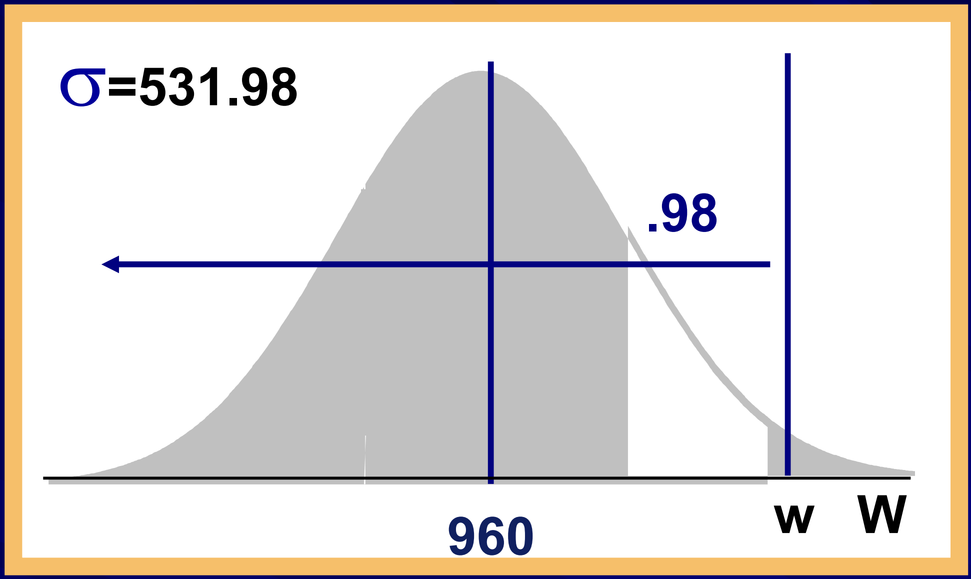
$Var(W) = Var(X) + Var(Y) + 2\sigma_X \sigma_Y Corr(X,Y)$

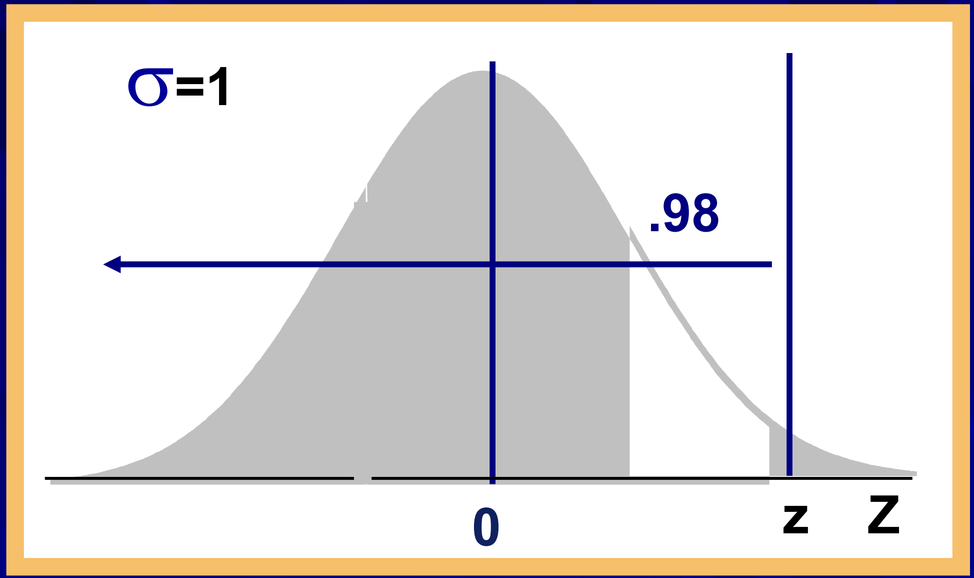
$= 250{,}000 + 10{,}000 + 2 * 500 * 100 * 0.23 = 283{,}000$

$\sigma_w = sqrt\{283{,}000\} = 531.98$ computers

- What is the 98th percentile of the demand for iMacs at the merged company?
  (In other words, find w such that $P(W \leq w) = 0.98$)

This will require reading the CDF table for Z backwards! (draw a picture!)





P(W≤w) = 0.98) (P(Z ≤z) = 0.98

z = (w – 960)/531.98.

Solving for w we get: w= 960 + 531.98z

Finding the z value from the table such that F(z) = 0.98, requires reading the table backwards….

F(2.05) = 0.9798 and F(2.06) = 0.9803, interpolating we see that z such that F(z) = 0.98 is approximately 2.054. Thus,

w= 960 + (2.054)531.98 = 2052.68. (The 98^{th} percentile is 2053 computers.)

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

Summary and Next Class

- The Normal distribution is very common
- Many real-world phenomena approximate the Normal distribution so we can use statistics about the Normal distribution to describe those phenomena
- Even phenomena that are distributed in other ways, e.g., lognormal, can be transformed into Normal and then analyzed easily
- As we will see in Lecture 10, all phenomena can be related to the Normal distribution using the Central Limit Theorem