AP Statistics Notes – Unit Two: The Normal Distributions

Syllabus Objectives: 1.5 – The student will summarize distributions of data measuring the position using quartiles, percentiles, and standardized scores (*z*-scores).

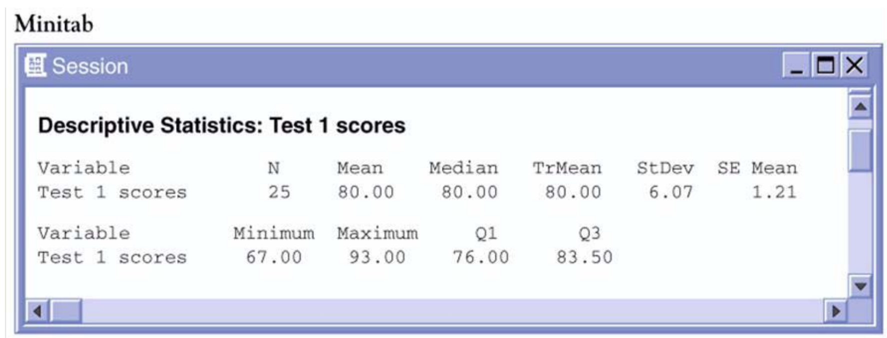
This unit introduces you to the concept of describing an observation's location within a distribution. You will learn how to use the Normal distribution to find standardized values, percentile ranks, and proportions of observations on intervals. You will also be introduced to methods for assessing the Normality of a distribution.

- Measuring relative standing: Consider the following test scores for a small class:

79	81	80	77	73	83	74	93	78	80	75	67	73
77	83	86	90	79	85	83	89	84	82	77	72	

Jenny's score is noted in red. How did she perform on this test relative to her peers?

6 | 7
7 | 2334
7 | 5777899
8 | 00123334
8 | 569
9 | 03


Minitab

Session

Descriptive Statistics: Test 1 scores

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Test 1 scores	25	80.00	80.00	80.00	6.07	1.21

Variable	Minimum	Maximum	Q1	Q3
Test 1 scores	67.00	93.00	76.00	83.50

The mean, shown above on the Minitab printout, is 80. Her score is "above average", but how far above average is it?

- One way to describe relative position in a data set is to tell how many standard deviations above or below the mean an observation is.
  - **Standardized Value: "z-score"**
    - If the mean and standard deviation of a distribution are known, the "z-score" can be found.
    - $z = \dfrac{x - \text{mean}}{\text{standard deviation}}$
    - A *z*-score tells us how many standard deviations away from the mean the observation falls.

Ex: From above, the mean test score is 80 and the standard deviation is 6.07 points.

To find Jenny's standardized *z*-score, $z = \dfrac{x-80}{6.07} = \dfrac{86-80}{6.07} = 0.99$.

Jenny's score is almost one full standard deviation above the mean. For a person scoring 80, they would have a *z*-score of 0, and for a person scoring below the mean, their *z*-score would be negative.

- Standardized values can be used to compare scores from two different distributions.
  - ✓ Statistics Test: mean = 80, std dev = 6.07
  - ✓ Chemistry Test: mean = 76, std dev = 4
  - ✓ Jenny got an 86 in Statistics and an 82 in Chemistry.
  - ✓ On which test did she perform better?

Statistics

$$z = \frac{86 - 80}{6.07} = 0.99$$

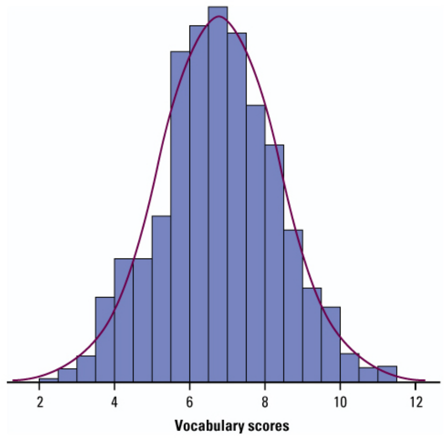
Chemistry

$$z = \frac{82 - 76}{4} = 01.50$$

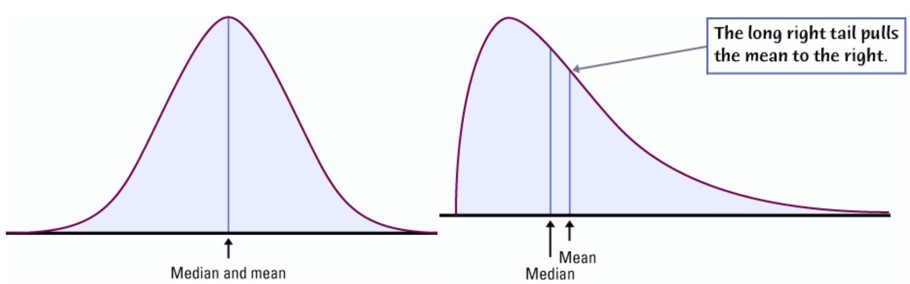
Although Jenny had a lower Chemistry score, she performed *relatively* better in Chemistry!

Syllabus Objectives: 1.11 – The student will compare distributions of data with respect to their shapes.

- In Unit 1, we learned how to plot a data set to describe its shape, center, spread and unusual features.
  - Sometimes, the overall pattern of a large number of observations is so regular that we can describe it using a smooth curve.
  - **Density Curve:** An idealized description of the overall pattern of a distribution.
    - Area underneath the curve = 1, representing 100% of the observations.
    -



    - Density curves come in many different shapes; symmetric, skewed, uniform, etc.
    - The area of a region of a density curve represents the % of observations that fall in that region.
    - The median of a density curve cuts the area in half.
    - The mean of a density curve is its "balance point."
      -

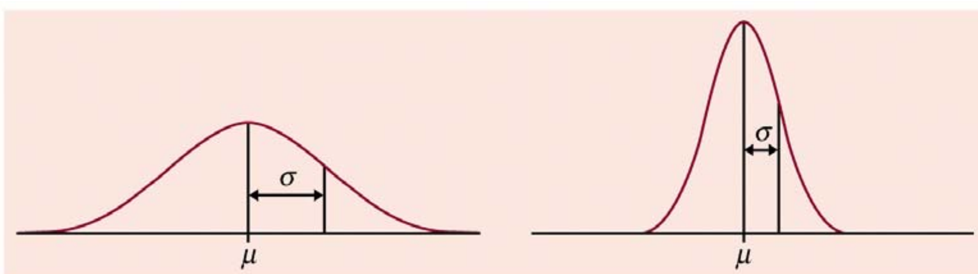


      - The median and mean are the same for a symmetric density curve. Both lie at the center. The mean of a skewed curve is pulled away from the median in the direction of the long tail.

Syllabus Objectives: 3.12 – The student will describe the properties of the Normal distribution.

- A very common type of density curve, known as the Normal distribution, is widely used.
  - These curves provide a reasonable approximation to the distribution of many different variables.
  - They play a central role in many of the inferential procedures discussed later.

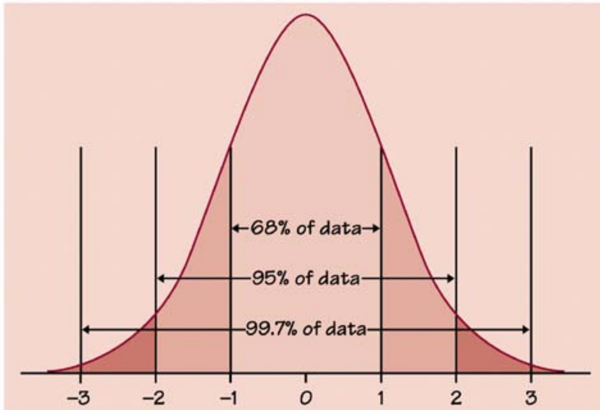
- Normal distributions are continuous distributions with the following properties:
  - The curve is single-peaked (unimodal).
  - The shape is symmetric.
  - More specifically, the distribution is bell-shaped.
  - The curve is described by two things: its mean, μ, and its standard deviation, σ. **Parameters** are *population* variables. They are represented by Greek letters, like μ and σ. **Statistics** are *sample* variables. They are represented by Roman letters, like $\overline{x}$ and *s*. A parameter describes a population characteristic and a statistic describes a sample characteristic.
  - Examples of Normal distributions:



This is more spread out. The distr. has a larger σ.

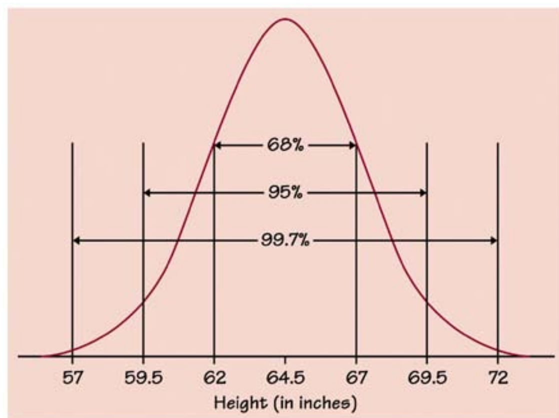
This is less spread out. The distr. has a smaller σ.

  - Inflection points – the points at which a change of curvature takes place. They are located at a distance σ on either side of μ.

- **The 68-95-99.7 Rule, also known as the Empirical Rule**
  - In a Normal distribution, 68% of the observations fall within one σ of μ.
  - 95% of the observations fall within two σ of μ.
  - 99.7% of the observations fall within three σ of μ.
  - 

  - Notation: $N(\mu,\sigma)$
  - If *N*(64.5,2.5), then we have a Normal distribution with a mean of 64.5 and a standard deviation of 2.5.

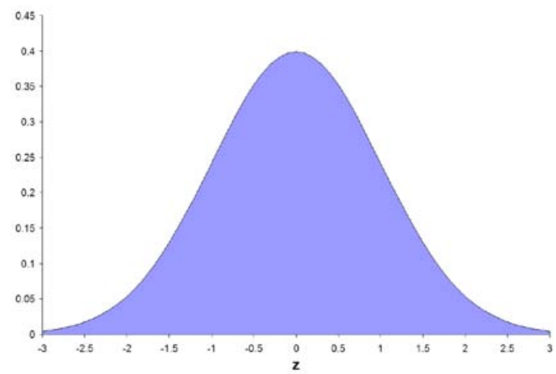
Here is the Normal distribution of heights with a mean of 64.5 inches and a S.D. of 2.5 inches. The 68-95-99.7 Rule shows that 68% of the heights will be found between 62 and 67 inches, 95% of the heights will be found between 59.5 and 69.5 inches and 99.7% of the heights will be between 57 and 72 inches.



- **Ex:** Suppose that we know that SAT I Math scores follow an approximately Normal distribution with mean 500 and standard deviation 100. *N*(500,100) Using the Empirical Rule, approximately 68% of students scored between 400 and 600 (within 1 S.D.), 95% scored between 300 and 700 (2 S.D.) and 99.7% scored between 200 and 800 (3 S.D.).

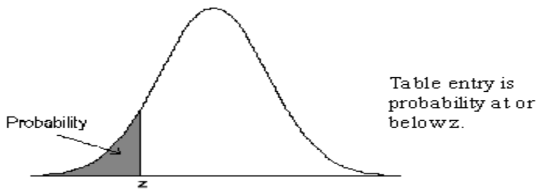
- **Ex:** Use the Empirical Rule to find percentiles. Since 500 is our mean and median (symmetric), then 500 is the 50^{th} percentile. 600 is one standard deviation above the mean. Because the curve is symmetric, then, 34% of the data lies between 500 and 600, so 600 is at the (50+34) 84^{th} percentile. 400 is one standard deviation below the mean and would be at the (50-34) 16^{th} percentile. Similarly, 700 is two standard deviations above the mean and using symmetry again, 47.5% of the data lies between 500 and 700, so 700 is the (50+47.5) 97^{th} percentile and 300 is the (50-47.5) 2.5^{th} percentile.

Syllabus Objectives: 3.13. – The student will solve problems using tables of the Normal distribution. 3.14 – The student will solve problems using the Normal distribution as a model for measurements.

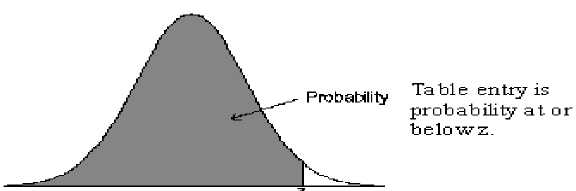
- **The Standard Normal Distribution**
  - A Normal distribution with a mean of 0 and a standard deviation of 1.
  - Notation: $\mathrm{N}(\mu,\sigma) \to \mathrm{N}(0,1)$
  - All Normal distributions can be standardized by using the *z*-score formula. This gives us a common scale to compute probabilities, which are areas under a Normal curve and above given intervals.
  - 

- **Normal Distribution Calculations**
  - Use the Empirical Rule if observations are 1, 2 or 3 standard deviations from the mean.
  - If we are unable to use the rule, we must standardize the distribution and use the Normal distribution table.
    - Step 1: Area under a density curve = proportion of the observations in the distribution.
    - Step 2: Standardize the distribution using the *z*-score formula.
    - Step 3: Look up the *z*-score in the Standard Normal Probabilities Table. It is a table of areas under the standard Normal curve. It reports the area under the curve from that value and BELOW. It gives the area to the LEFT of the score.
  - Using the *z* table
    - Find the *z* score for the problem using the formula and round to the nearest hundredth.
    - Look it up on the table. If you want the proportion to the left of that number, your answer is the 4-digit number you find. If you want the area to the right of that *z*-score, subtract the 4-digit number from one. Remember, the curve is symmetric, so if you want the area to the right of z = 1.23, that is the same as the area to the left of z = -1.23.
    - The Normal distribution is describing **continuous** data. This means that the proportion of observations with $x > a$ is the same as the proportion with $x \geq a$. There is no area above a single point.
    - To find the area between two scores, find the *z* score for each value, look both of them up on the table and subtract the two areas to find the area BETWEEN the two scores.

- Normal Tables:



z*	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-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	0.0010	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	0.0019	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	0.0046	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0061	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
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
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

- Table Example: Find $P(z < 0.46)$.

Column labeled 0.06

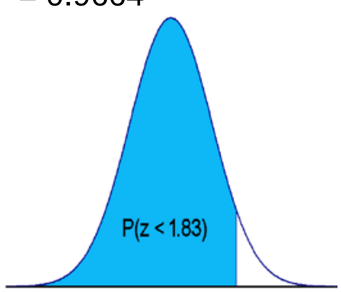
z*	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293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

Row labeled 0.4

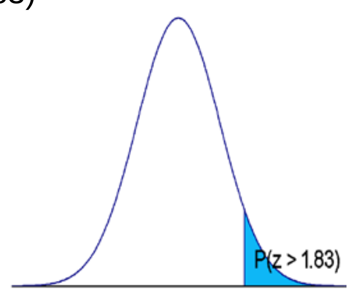
P(z < 0.46) = 0.6772

- Table Example 2:

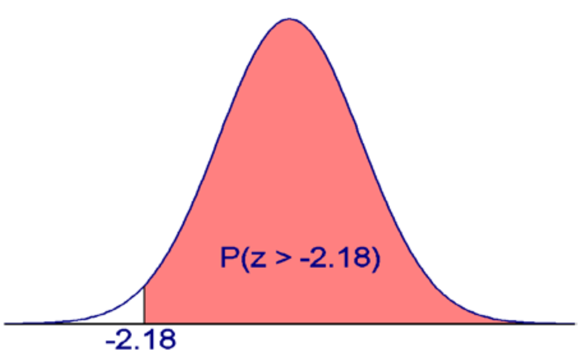
(a) P(z < 1.83)
= 0.9664

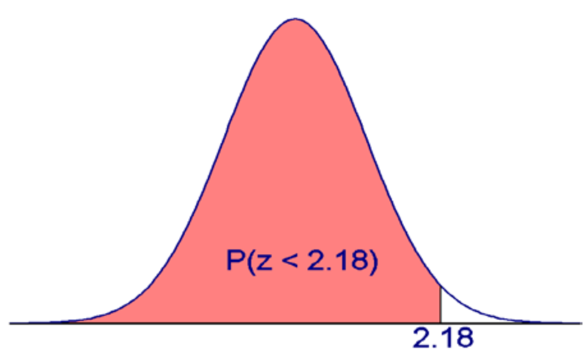


(b) P(z > 1.83)
= 1 – P(z < 1.83)
= 1 – 0.9664
= 0.0336



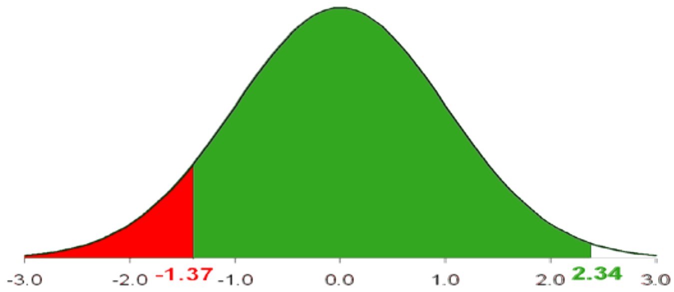
- Symmetry Property: $P(z > z^*) = P(z < -z^*)$





P(z > -2.18) = P(z < 2.18) = 0.9854

- Table Example 3: Find P(-1.37 < *z* < 2.34).

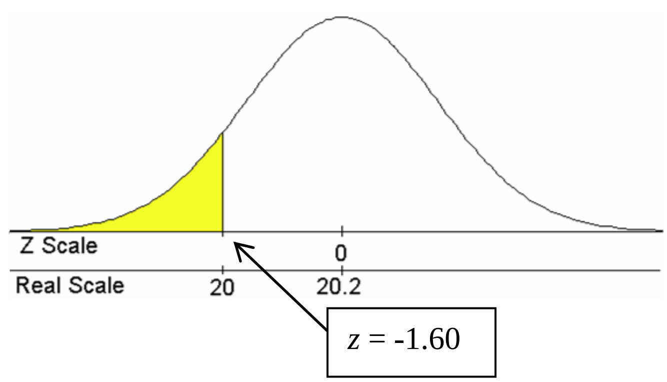


P(Z<2.34)=0.9904

P(Z<-1.37)=0.0853

$$P(-1.37 < z < 2.34) = 0.9904 - 0.0853 = 0.9051$$

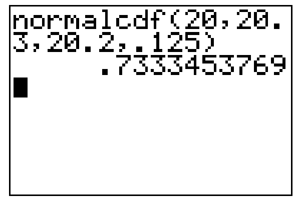
- Finding Normal Probabilities
    - Example 1: A company produces "20 ounce" jars of picante sauce. The true amounts of sauce in the jars of this brand sauce follow a Normal distribution. The contents of the jars are Normally distributed with a true mean of $\mu = 20.2$ ounces and a standard deviation of $\sigma = 0.125$ ounces. What proportion of the jars are under-filled? (have less than 20 ounces of sauce).
        - Step 1: Write the probability statement and standardize the score. $P(x < 20) = \frac{20 - 20.2}{0.125} = \frac{-0.2}{0.125} = -1.60$. We have standardized the distribution. Therefore, we need to find the $P(z < -1.60)$.
        - Step 2: Draw the Normal curve, shade the appropriate area.



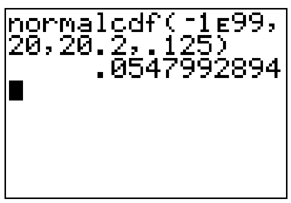
        - Step 3: Look up the value of -1.60 on the *z* table and find the value 0.0548.
        - Step 4: Summarize: That means, the proportion of the sauce jars that are under-filled is 0.0548.
    - Example 2: What proportion of the sauce jars contain between 20 and 20.3 ounces of sauce.
        - Step 1: Find BOTH standardized scores. We have the one for 20, standardize 20.3: $P(20 < x < 20.3)$. $\frac{20.3 - 20.2}{0.125} = \frac{0.1}{0.125} = 0.80$.

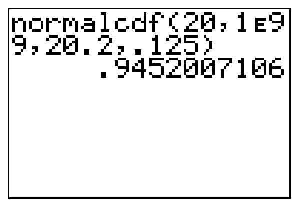
- Step 2: Find $P(-1.60 < z < 0.80)$.
- Step 3: Look up both *z* values on the table and subtract the two probabilities: $0.7881 - 0.0548 = 0.7333$.
- Step 4: Summarize: The proportion of sauce jars between 20 and 20.3 ounces is 0.7333.

- **Normal probabilities on the TI-84 Graphing Calculator**
  - The normalcdf command can be used to find areas under a Normal curve.
  - Press [2nd] [VARS] (DISTR) and choose $2:\text{normalcdf}($.
  - Complete the command. It is waiting for four values, [lower limit, upper limit, mean, standard deviation]. Use 1E99 for values to the right end of the curve, and -1E99 for the left end of the curve. Example 2 $\Rightarrow$

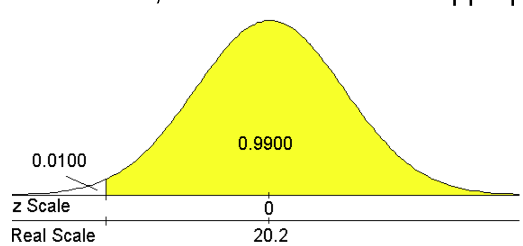


  - Finding $P(x < 20)$ and $P(x > 20)$. {Note: probabilities add up to 1!}





- **Normal Calculations – Finding a value, given a proportion**
  - Now find the *z*-score if given a proportion area. We use the table backwards.
    - Step 1: Look in the MIDDLE of the table for the CLOSEST decimal related to the percent or portion given in the problem.
    - Step 2: Read OUT along the table's row and column to find the *z*-score.
    - Example: $P(z < 0.44) = .67$, so 67% of all *z* values are less than 0.44, and 0.44 is the 67th percentile of the standard normal distribution.
  - Example 2 (from above) continued: 99% of the jars of this brand of picante sauce will contain more than what amount of sauce? Notice a *proportion* is given in the problem and we will find the *x* value that corresponds to 0.99.
    - Step 1: Draw the curve, label and shade the appropriate area.



    - Step 2: Recall that the z table shows areas to the LEFT of the z value, so subtract (1-0.9900) and look up 0.0100 on the *z* table.

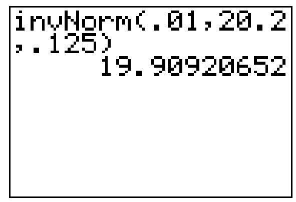
z	0.00	0.01	0.02	0.03	0.04	0.05
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071
-2.3	0.0107	0.0105	0.0102	0.0099	0.0096	0.0094
-2.2	0.0139	0.0135	0.0132	0.0129	0.0125	0.0122

When we look up 0.0100 in the BODY of the table, we do not find this value. The entry closest to 0.0100 is 0.0099 corresponding to the *z* value -2.33.

- Step 3: Use the standardized formula to solve for *x*. Our formula: $z = \dfrac{x-\mu}{\sigma}$ becomes $x = z \cdot \sigma + \mu$.
  Solving: $x = (-2.33)(0.125) + 20.2 = 19.91$
- Step 4: Summarize – 99% of the jars of picante sauce of this brand will contain more than 19.91 ounces.

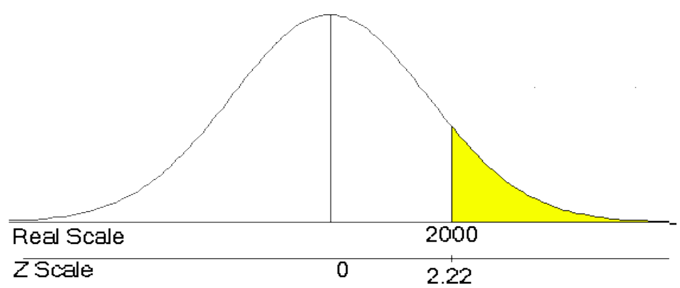
- **Finding values with invNorm on the TI-84 Graphing Calculator**
  - The invNorm function calculates the raw or standardized Normal value corresponding to a known area under a Normal curve.
  - Press 

(DISTR) and choose 3:invNorm(.
  - Complete the command. It is waiting for three values. Input the proportion as a decimal, the mean and the standard deviation. Example 2 continued ⇒



You try:

- Final Normal probability example: The time to first failure of a unit of a brand of ink jet printer is approximately Normally distributed with a mean of 1,500 hours and a standard deviation of 225 hours.
  - (a) What proportion of these printers will longer than 2,000 hours?
    - Step 1: $P(x > 2000) = P\left(z > \dfrac{2000-1500}{225}\right) = 2.22$
    - Step 2:

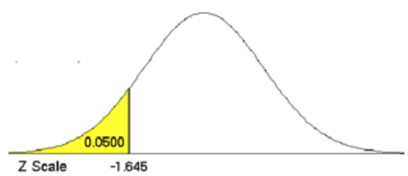


    - Step 3: Find $P(z > 2.22)$. Look up 2.22 and calculate. $P(z > 2.22) = 1 - 0.9868 = 0.0132$.
    - Step 4: The proportion of printers that will last 2,000 hours or more is 0.0132.
  - (b) What proportion of these printers will last between 1,300 and 1,800 hours?

- Step 1: $P(1300 < x < 1800)$ $z = \dfrac{1300 - 1500}{225} = -0.89$ and $z = \dfrac{1800 - 1500}{225} = 1.33$. $P(1300 < x < 1800) = P(-0.89 < z < 1.33)$.
- Step 3: Look up -0.89 and 1.33 and find the difference. $0.9082 - 0.1867 = 0.7215$
- Step 4: The proportion of printers lasting between 1,300 and 1,800 hours is 0.7215.

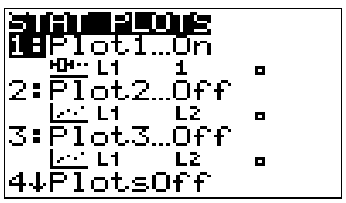
▪ (c) What should be the guarantee time for these printers if the manufacturer wants only 5% to fail within the guarantee period?

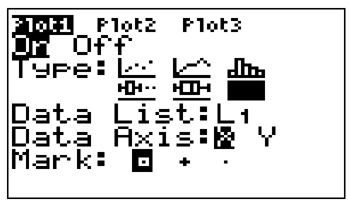
- Step 1: Interested in the lowest 5%.
- Step 2: Look up 0.0500 in the body of the normal table. It is not there, but there are two values that are close: 0.0505 corresponding to a *z*-score of -1.64 and 0.0495 corresponding to a *z*-score of -1.65. Since 0.0500 is right in between these two values, the *z* value we want is -1.645.

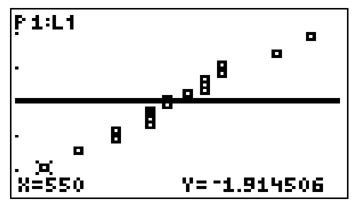


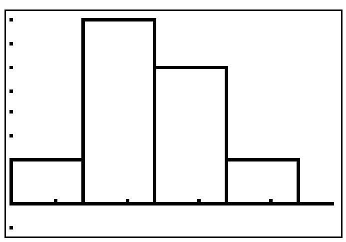
- Step 3: $x = z \cdot \sigma + \mu \Rightarrow x = (-1.645)(225) + 1500 = 1129.875$.
- Step 4: The guarantee period should be 1130 hours.

- Is the distribution Normal?
  - In the previous examples, we've assumed that the underlying data distribution was roughly unimodal and symmetric. One must CHECK to see whether the Normal model is reasonable.
    - Plot a histogram, stemplot and/or boxplot to determine if a distribution is bell-shaped.
    - Determine the proportion of observations within one, two, and three standard deviations of the mean, and compare with the 68-95-99.7 rule for Normal distributions.
    - Construct and interpret a Normal probability plot.
      - The **Normal probability plot** plots each data point *x* against its corresponding *z* score. The *x*-values are normally graphed on the horizontal axis and the *z*-scores on the vertical axis.
      - If the distribution of the data is roughly Normal, the plot is roughly a diagonal straight line. Deviations from the straight line indicate it is not Normal.
      - Graphing Calculator Example – data below is 19 breaking strengths of connections of fine wires to semiconductor wafers (in pounds):

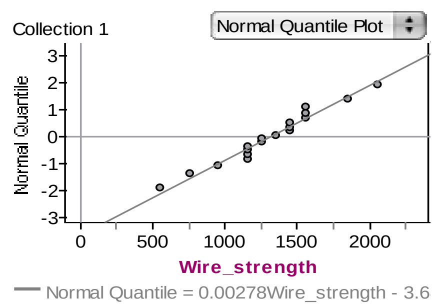






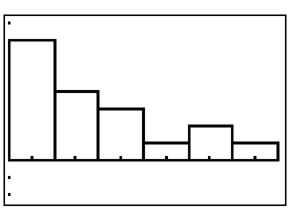


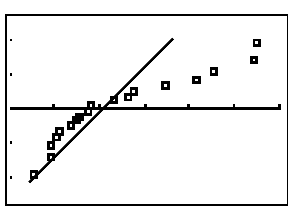
Although the data is not perfectly Normal, the histogram shows a unimodal, fairly symmetric distribution.

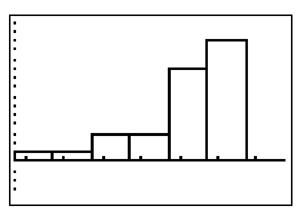


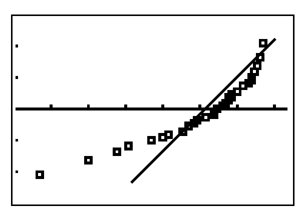
Normal Quartile (Probability) Plot in Fathom. Note: All *z*-scores are between -2 and 2. The Empirical Rule states, 95% of the data should be within 2 s.d. of the mean.

      - Skewed distributions and their Normal probability plots.









Skewed to the right – Normal probability plot "falls off" on the right.

Skewed to the left – Normal probability plot "falls off" on the left.