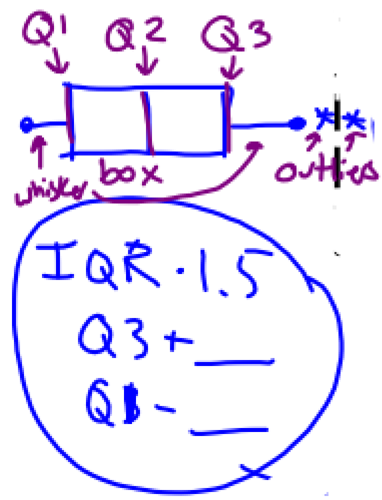
Station 1 – Measures of Central Tendency and Variation

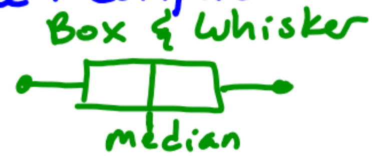
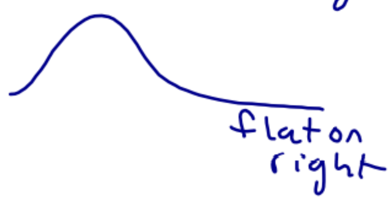
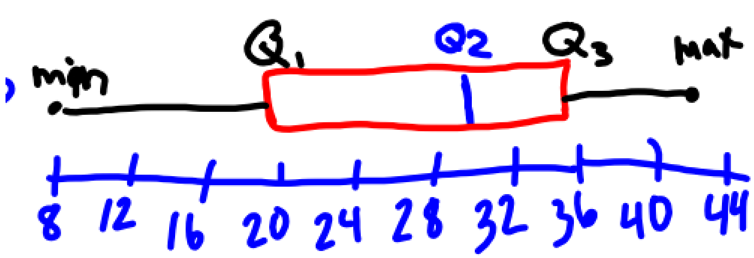
Definition Central Tendency	**<u>How should we measure the Center? Where does the Center tend to be?</u>** **<u>Mean</u> –** the Center or Average. Add up every piece of data and divide by how many items there were. Outliers mess with the mean and drag the center towards them. **<u>Median</u>** – the Center or Middle data item. Order the data from least to greatest, including repeats, and find the exact middle. If there are two numbers in the middle, take the average (mean) of those two numbers (add the two numbers and divide by 2). Outliers do not mess with the median, they are tossed aside as the center is found. **<u>Mode</u>** – The most repeated number. If the amount of times a number is repeated is the same as for another number, then the data is considered Bi-modal and has two modes. Data could have more than one or two modes. Outliers do not effect the mode.
Definition Variation	**<u>How should be measure or display the spread of data? How much or how wide does the data vary?</u>** **<u>Range</u>** – The distance between the maximum and the minimum. Subtract the biggest and the smallest number. Maximum – the biggest data value Minimum – the smallest data value **<u>IQR</u>** – Inter Quartile Range (see Box and Whisker Plots) the distance between the first and third quartile. The breadth of the box part. Subtract Q3 – Q1 Q3 is the upper quartile, the median of the upper half of data Q1 is the lower quartile, the median of the lower half of data **<u>Standard Deviation</u>** – the distance from the Mean that can give a probability percent for the likely hood of randomly getting that data value.

Station 2 – Box and Whisker Plots.

<table>
<tr><td>Definition</td><td colspan="2"><u>A data display that depicts the spread of data using the Medians.</u>
The box part is 50% of the data and each whisker is 25%
The Box is created by the Median, and the upper half median and the lower half median.
You can clearly see outliers (data points that are more like flukes than reasonable)</td></tr>
<tr><td>Steps for Calculating and Drawing The Box</td><td>1) Find the Median or Q2: List all data values in order from least to greatest and find the median and draw a vertical line
2) Find the Lower Quartile or Q1: Find the median of the lower half set of data and draw a vertical line
3) Find the Upper Quartile or Q3: Find the median of the upper half set of data draw a vertical line
4) Finish drawing a box around these quartiles
5) Calculate Outliers: Find IQR (Inter Quartile Range) (subtract Q3 - Q1) and multiply by 1.5. Add to Q3 and subtract fromQ1</td><td>
</td></tr>
<tr><td>How to Calculate Outliers</td><td colspan="2">How far out is too far out? Any thing farther than the distance of the box and a half above or below the box is too far.

1. Take Q3 – Q1
2. Multiply this IQR by 1.5 to get the Outlier Distance
3. Add the Outlier Distance to Q3, anything bigger than that value is an Outlier
4. Subtract the Outlier distance from Q1, anything smaller than this is an outlier.

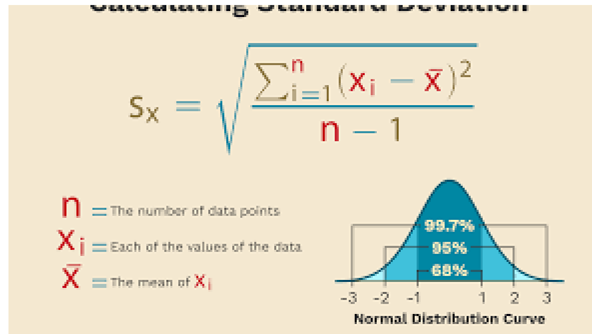
IQR: Q3-Q1
17-4=13
Outlier Distance: IQR·1.5
13·1.5=19.5
• Q3+19.5 = any thing above
17+19.5=36.5 is an outlier
45 is an outlier
• Q1-19.5
4-19.5=-15.5 any thing below is an outlier
no bottom outlier</td></tr>
</table>

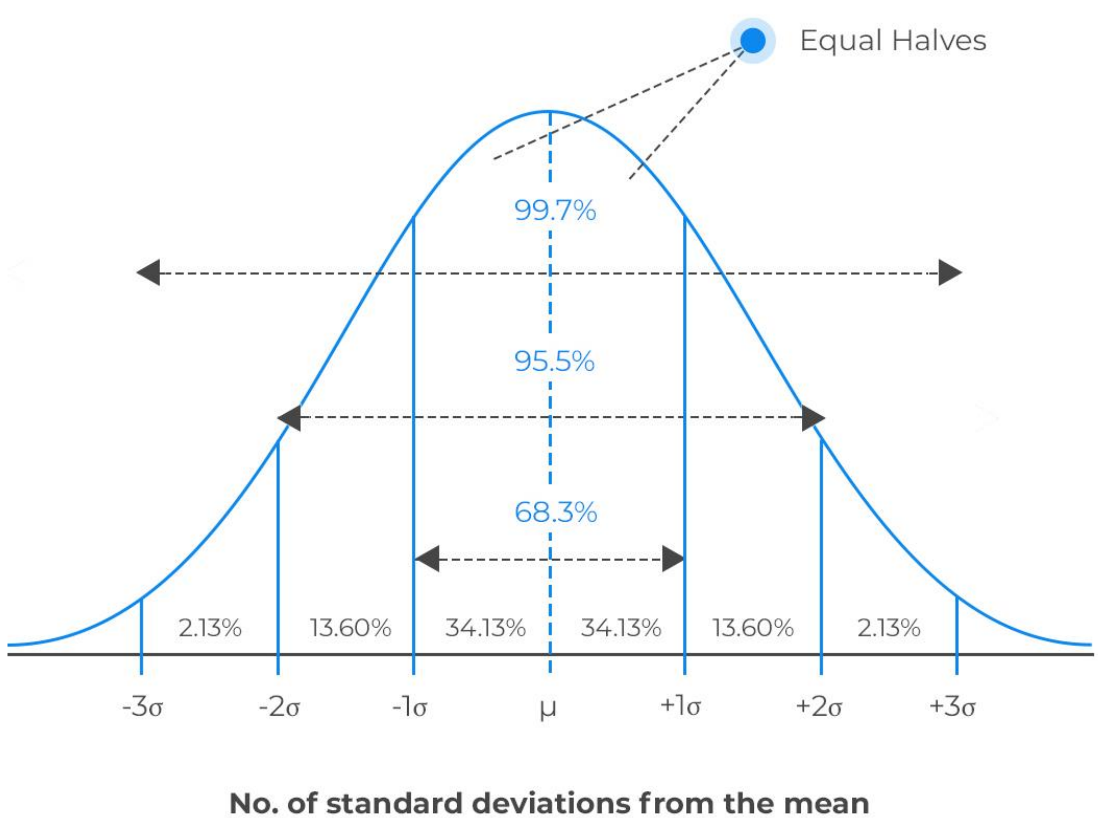
Spread of Data	**Distribution** the spread of data far distance : spread out close distance : compact   **Skewed Left** Screwed up on the left (longer tail on the left)   **Skewed Right** Screwed up on the right (longer tail on the right)  
Extra Information	Box and Whisker plots break up the data set into 4 parts, where the same amount of data points are within each Quartile You can clearly see how close or spread out each quarter of the data points were
Example	40 Students were surveyed how long they study for their math tests.  How many students studied less than half an hour? *50% of the data is below Q2 at 30, so 50% of 40 people is 20 people, 20 people studied less than 30 minutes* What percentage of people study between 20 and 35 minutes? *50%, which would be 20 people.* What percentage of people study more than 35 minutes? *25%, which would be 10 people*

Station 3 – Normal Distribution

<table>
<tr><td>Definition</td><td colspan="2"><u>A bell curved shape is made when the Mean, Median, and Mode are all the same</u></td></tr>
<tr><td>How to label a Normal Curve</td><td>$\bar{x}$ or μ -

The middle of the curve is the Mean or average of the data

s or σ -
The Standard Deviation is how much away the data piece has deviated from the Mean or center of the data. The bigger the standard deviation, the more spread out the data is.
Add the s or σ to the Mean to get the 1st standard deviation away, add it again, to get the 2nd standard deviation away, and it again to get the 3rd standard deviation away. Repeat with subtraction.</td><td>Calculate Standard Deviation using the following formula:


Take each data piece, and subtract the Mean from it. Square that difference. Add it to the next squared difference from the next data piece. Repeat. Divide by 1 less than the total number of data pieces. Square root the result.</td></tr>
<tr><td>Diagram</td><td colspan="2">

The area under the curve (the percentages) between values is the probability that the data falls in between those values</td></tr>
</table>

Population Vs Sample

We distinguish between population and sample characteristics by referring to **population** characteristics as **population parameters** and **sample** characteristics as **sample statistics**. Thus, the **mean** of a **population** is a **parameter** of that **population**, but the **mean** of a **sample** is a **statistic** of that **sample**. Beyond that distinction there is no difference in the naming or computing for the measures **mode**, the **median**, the **range**, and the **quartiles**. The **mean of a population, μ**, has a different symbol from the one used for the **mean of a sample,** $\overline{x}$, however the computation of each is the same. **Standard deviation** has both different symbols, **σ** for a **population** and s_x for a **sample**, and slightly a slightly different formula.

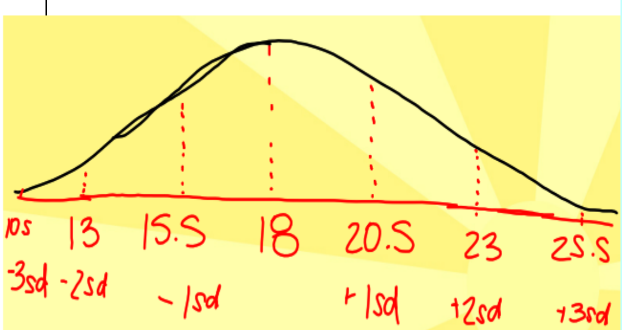
Extra Information

Table 1: Mean and Standard Deviation

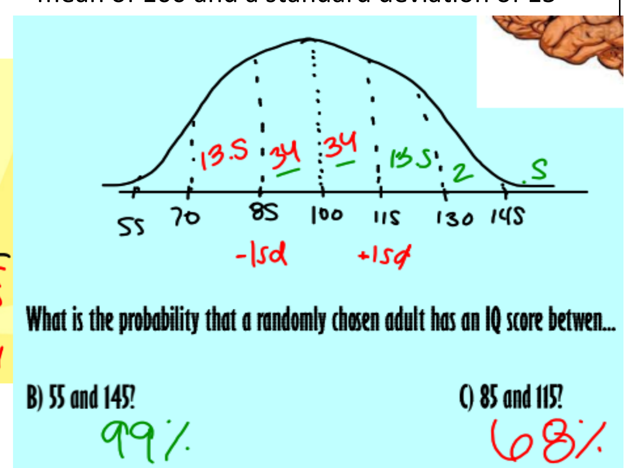
Measure Name	Symbol for Population	Symbol for Sample	Computation for Population	Computation for Sample
Mean	μ	$\overline{x}$	$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$	$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Standard Deviation	σ	s_x	$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$ or the equivalent form $\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x^2) - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}}{N}}$	$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}}$ or the equivalent form $s_x = \sqrt{\frac{\sum_{i=1}^{n}(x^2) - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}}$

Examples Normal Curve

Normal Curve with a mean of 18 and a standard deviation of 2.5



Adult IQ scores are normally distributed with a mean of 100 and a standard deviation of 15



Standard Deviation

Data sample: 5,5,7,7,11,14,14,14,18, 20,22 Mean is 13.7

Standard Deviation = $\sqrt{\frac{(5-13.7)^2+(5-13.7)^2+(7-13.7)^2+(11-13.7)^2+(14-13.7)^2+(14-13.7)^2+(5-13.7)^2+\cdots}{10-1}}$

or $= \sqrt{\frac{2(5-13.7)^2+2(7-13.7)^2+(11-13.7)^2+3(14-13.7)^2+(18-13.7)^2+(20-13.7)^2+(22-13.7)^2}{10-1}}$

$= \sqrt{\frac{2(-8.7)^2+2(-6.7)^2+(-2.7)^2+3(0.3)^2+(4.3)^2+(6.3)^2+(8.3)^2}{9}} = \sqrt{\frac{375.79}{9}} =$

$\sqrt{41.75444\ldots}$ = 6.462 is standard deviation amount you add and subtract from the mean.

Station 4 – Z-Score

z-score

How many standard deviations a data value is above or below the mean

Positive z-score is above the mean

Negative z-score is below the mean

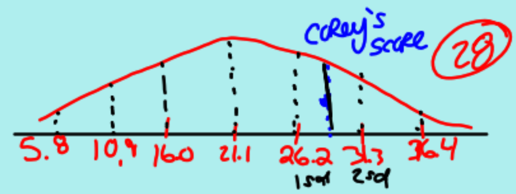
$$z - score = \frac{Data\ Value - Mean}{Standard\ Deviation}$$

Example 2) Determine the z-scores for Corey's ACT score and Danielle's SAT score to verify who performed better on their respective exam.

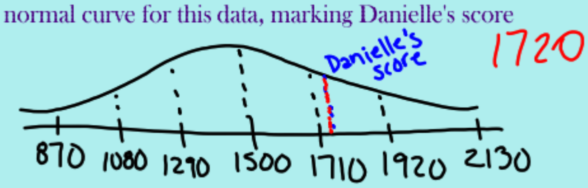
Corey's z-score: $\frac{28 - 21.1}{5.1} = 1.35$ standard deviations above average

Danielle's z-score: $\frac{1720 - 1500}{210} = 1.05$

A) The average ACT score is 21.1 with a standard deviation of 5.1 Sketch the normal curve for this data, marking Corey's score as well



B) The average SAT score is 1500 with a standard deviation of 210 Sketch the normal curve for this data, marking Danielle's score



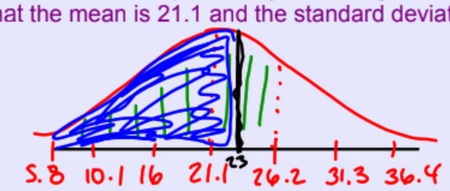
Z-Score Table

Steps for finding probability of a Normal Distribution when numbers don't fall nicely on Standard Deviations:

1) Find the z-score of the data value given

2) Use the z-score table provided to determine the probability of anything LESS THAN that value

Example 3) What is the probability of scoring less than a 23 on the ACT, given that the mean is 21.1 and the standard deviation is 5.1



z-score: $\frac{data - mean}{s.d.}$

$\frac{23 - 21.1}{5.1} = .37$

$= .6443$ or 64.43%

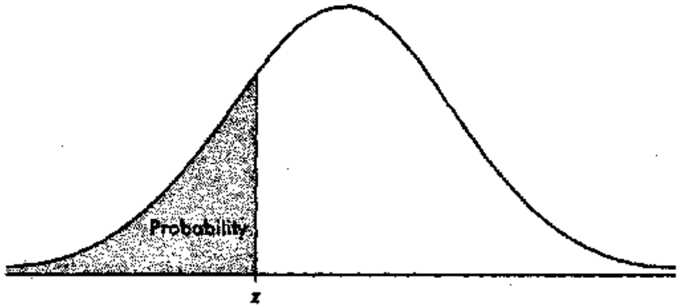


TABLE A: STANDARD NORMAL PROBABILITIES

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

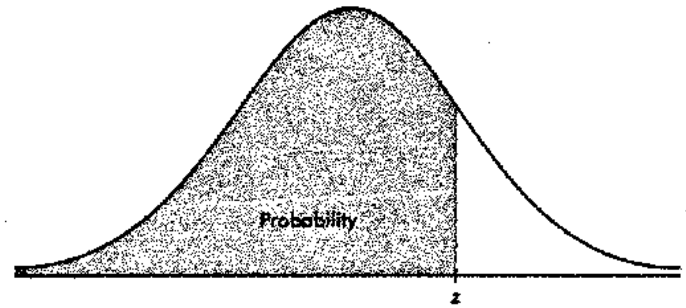


TABLE A: STANDARD NORMAL PROBABILITIES (CONTINUED)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998