12_normalcurve.pdf	Michael Hallstone, Ph.D. hallston@hawaii.edu

Lecture 12: Normal Probability Distribution or "Normal Curve"

The real importance of this lecture is to show you what a normal curve looks like (it looks like a "bell curve"), to show you that area under the curve can be used to determine probability, and to make you appreciate that the z-score formula saves you a lot of tedious algebra rules!

Start in person class with coin toss [online students can ignore this section]

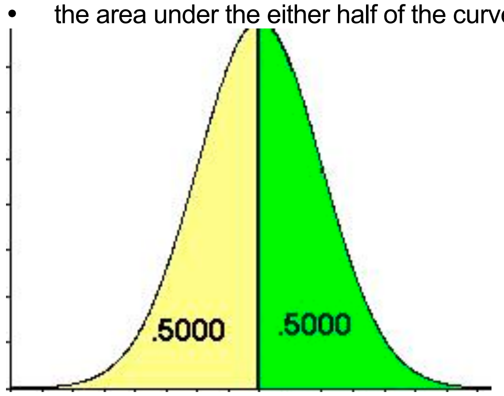
Have students toss a coin 10 times and count the number of heads. Plot the number of heads 1-10 on the x axis and frequency on the y axis. Have them do it 10 more times, plot. Repeat as necessary.

Introduction to the Normal Distribution (Bell Curve)

Technically speaking the normal or "gaussian" distribution relates to a probability distribution for a continuous variable, not a coin toss [which is discrete]. But hopefully the coin exercise started to resemble a normal curve.

Characteristics of the Normal Curve

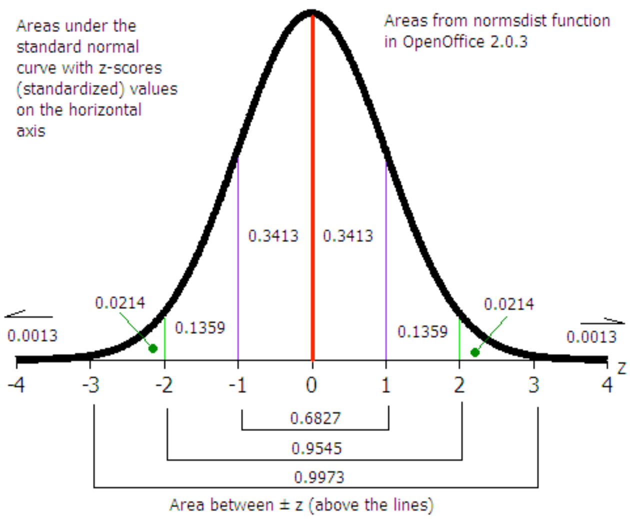
- symmetrical sides – mirror images of each other mean=median=mode
- tails go out to infinity
- bell shaped and only differs based upon dispersion
- not a single curve but it is family of curves ("n" stay tuned!)
- area under =1 or 100% related to probability
- the area under the either half of the curve is 50% -- add both halves up and it's 100%



Source: jamesstack.com

Area Under the Normal Curve Using Standard Deviations

Using the normal curve, probability is determined by the area under curve between two points. We cannot know probability of exact score but we can know probability of a score falling within a range of values. Below is a picture of a normal curve with some probabilities associated with the normal curve. The numbers at the bottom are standard deviation units. So for example 34.13% (.3413) of the area under the normal curve falls between 0 and 1 standard deviations. With a normal curve the probability of a score falling between -1 and +1 standard deviations is 34.13% + 34.13% = 68.26%. So, the scores on the SAT test are normally distributed. That means if you select a person at random there is a 68.26% chance that they scored within 1 and +1 standard deviations from the mean. Or



68.26% of the people who took the SAT scored between 1 and +1 standard deviations from the mean.

Online students can ignore this: Note to Mike from Mike: for an in person class Draw 2 pictures: 1) 3 curves with same mean but diff stand div 2) 3 curves with same SD but diff means.

Introducing z-scores

What do we do when we want to compare scores with different means and standard distributions? So pretend you work for an organization that tests all new hires in 4 areas: computers, writing, management, and customer service. The scores on these tests measure a person's skills in each of these areas. For example if a person scores really well on computers, but not so well in customer service, it would be best to place them in a job where they work with computers rather than customers right?

How to we compare a person's scores on each of the tests? We have some rules to help us. They involve algebra. [Don't panic! Actually you won't have to do all this algebra! You will just use the z-score formula we will see in lecture 13: z-scores]

lets make sd's the same										
Area	test score	mean	SD		test score	mean	SD	new score	new mean	new sd
management	90	90	5		90x2	90x2	5x2	180	180	10
service	75	80	10		75x1	80x1	10x1	75	80	10
computers	60	50	5		60x2	50x2	5x2	120	100	10
writing	81	76	10		81x1	76x1	10x1	81	76	10
now need to make means the same (100)-- can add and subtract means and SD stays the same										
Area	new score	New mean	new sd		new score	new mean	sd stays	new score	new mean	new sd
management	180	180	10		180-80	180-80	10	100	100	10
service	75	80	10		75+20	80+20	10	95	100	10
computers	120	100	10		120	100	10	120	100	10
writing	81	76	10		81+24	76+24	10	105	100	10

After all that math, we are able to compare all scores where the mean and the SD is the same. Thus this person's test scores have the same mean and SD and can now be compared as seen in the highlighted column in the above table.

That is sort of hard to see, so here are the scores in a new table. See below

Area	new score
management	100
service	95
computers	120
writing	105

So putting them in order (as seen in the table below) this person's skills upon entering the organization appear to be strongest in computers, followed in order by writing, management, and customer service.

Area	new score
computers	120
writing	105
management	100
service	95

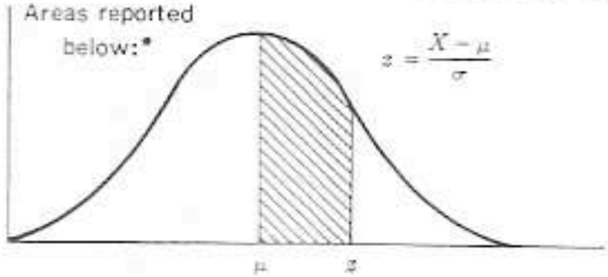
Thus the organization would be best served putting this person in job involving computers.

Practice

There really is not anything to practice for this lecture. The purpose is to introduce you to the normal curve and get you ready for z-scores and eventually probabilities associated with the normal curve. You don't have to do that algebra about the computers test in this class so don't panic.

Z table

AREAS UNDER THE STANDARD NORMAL PROBABILITY DISTRIBUTION



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4014
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4983	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4989	.4989	.4989	.4989	.4990	.4990
3.5	.4997									
4.0	.4999683									

* Example: For $z = 1.96$, the shaded area is 0.4750 out of the total area of 1.0000.