Technical Assistance Document

2009 Algebra II Standard of Learning AII.11

Acknowledgements

The Virginia Department of Education wishes to express sincere thanks to Gloria Barrett, J. Patrick Lintner, Donna Meeks, Stephanie Shea, Amy Siepka, and Michael Traylor for their assistance in the development of this document.

Questions about this document may be addressed to Michael Bolling at Michael.Bolling@doe.virginia.gov or (804) 786-6418.

Virginia Department of Education October 2010

Standard of Learning AII.11

The student will identify properties of a normal distribution and apply those properties to determine probabilities associated with areas under the standard normal curve.

Background

In Algebra I SOL A.9, students study descriptive statistics through exploration of standard deviation, mean absolute deviation, and z-scores of data sets. Students compute these values for small defined populations, and the focus of instruction is on interpretation of these descriptive statistics. Algebra II.11 continues the study of descriptive statistics as students analyze properties of normal distributions and apply those properties to determine probabilities associated with areas under the standard normal curve. Students will be provided with the mean, standard deviation, and/or elements of samples or populations of normal distributions and asked to apply the properties of normal distributions to calculate probabilities associated with given elements of the data set.

Normal distributions

Students examine many sets of data in their study of statistics. Some data have been provided to them and some data they have collected through surveys, experiments, or observations. Representation of data can take on many forms (line plots, stem-and-leaf plots, box-and-whisker plots, bar graphs, circle graphs, and histograms). Often teachers try to describe the physical shape of these representations in words or by using measures such as the arithmetic mean, the balance point of the data, or the median, the point at which the data is split into two equal amounts of data points. Another useful type of graphical representation is called a density curve, which models the pattern of a distribution. The distribution of certain types of data take on the appearance of a bell-shaped density curve called a normal curve. These collections of data are often naturally-occurring data or data produced by repetition in a mechanical process.

Examples of data that can be modeled by normal distributions include:

- heights of corn stalks in similar growing environments;
- heights of 16-year-old girls;
- blood pressures of 18-year-old males;
- weights of pennies in a given production year;
- lifespan of a specific electric motor; and
- standardized test scores like the ACT[®] or SAT[®].



Properties of normal distributions

Normal distributions are represented by a family of symmetric, bell-shaped curves called "normal" curves. Normal curves are defined by the mean and standard deviation of the data set. The arithmetic mean is located on the line of symmetry of the curve. In a normal distribution, the arithmetic mean is equivalent to the median and mode of the data set. Approximately 68 percent of the data values fall within one standard

Standard Deviation

The standard deviation (σ) of a data set is a measure of the spread of data about the mean. The greater the value of the standard deviation, the more spread out the data are about the mean. The lesser (closer to 0) the value of the standard deviation, the closer the data are clustered about the mean.

deviation (σ) of the mean (μ), approximately 95 percent of the data values fall within two standard deviations of the mean, and approximately 99.7 percent of the data values fall within three standard deviations of the mean. This is often referred to as the 68-95-99.7 rule.



Figure 1 shows the approximate percentage of observations that fall within different partitions of the normal distribution.

In normal distributions, the total area under the curve is always equal to 1. The mean and standard deviation of a normal distribution affect the location and shape of the curve. The vertical line of symmetry of the normal distribution falls at the mean and the width or spread of the curve is determined by the standard deviation. The greater the standard deviation, the wider ("flatter" or "less peaked") the distribution of the data.





Figure 2 shows how mean (μ) and standard deviation (σ) affect the graph of the normal distribution.

2

A normal distribution with $\mu = 5$ and $\sigma = 1.5$ can be graphed on a graphing calculator.



Determining probabilities associated with normal distributions

The cumulative probability of a specified range of values can be represented as the area under a normal distribution curve between the lower and upper bounds. When the range of values is an interval with lower and upper bounds equal to the mean or mean plus or minus one, two, or three standard deviations, the 68-95-99.7 rule can be used to determine probabilities. Graphing calculators can also be used to compute and graph the areas under normal curves.

Example 1

Given a normally distributed data set of 500 observations measuring tree heights in a forest, what is the approximate number of observations that fall within two standard deviations from the mean?

Solution

A quick sketch of the normal distribution will assist in solving this problem (refer to Figure 1). We know from the 68-95-99.7 rule that 95 percent of the data falls within two standard deviations from the mean. Therefore, approximately $500 \cdot 0.95 = 475$ of the trees' height observations fall within two standard deviations from the mean.

Example 2

A normally distributed data set containing the number of ball bearings produced during a specified interval of time has a mean of 150 and a standard deviation of 10. What percentage of the observed values fall between 140 and 160?

Solution

From the sketch of a normally distributed data set, 140 is one standard deviation below the mean while 160 is one standard deviation above the mean. Therefore, approximately 34.1% + 34.1% = 68.2% of the data in this distribution falls between 140 and 160.



Example 3

Donna's boss asked her to purchase a large number of 20-watt florescent light bulbs for their company. She has narrowed her search to two companies offering 20-watt bulbs for the same price. The Bulb Emporium and Lights-R-Us each claim that the mean lifespan for their 20-watt bulbs is 10,000 hours. The lifespan of light bulbs has a distribution that is approximately normal. The Bulb Emporium's distribution of the lifespan for 20-watt bulbs has a standard deviation of 1,000 hours and Lights-R-Us' distribution of the lifespan of 20-watt bulbs has a standard deviation of 750 hours. Donna's boss asked her to use probabilities associated with these normal distributions to make a purchasing decision.

Donna decided that she would compare the proportion of light bulbs from each company that would be expected to last for different intervals of time. She started with calculating the probability that a light bulb would be expected to last less than or equal to 9,000 hours. Letting *x* represent the lifespan of a light bulb, $P(x \le 9,000 \text{ hours})$ represents the probability that the lifespan of a light bulb would fall less than or equal to 9,000 hours in its normal distribution. Donna continued by finding $P(9,000 \le x \le 11,000 \text{ hours})$ and $P(x \ge 11,000 \text{ hours})$ for each company.

There are two ways to find the cumulative probability within a range of values on TI graphing calculators. The probability can be computed by finding the area under the curve bounded by a range of values using the ShadeNorm function, or it can be directly computed using the normalcdf function.



On Casio graphing calculators, a feature similar to the TI's ShadeNorm is only available in terms of the standard normal distribution using z-scores (see Example 4, Solution B on page 9). The cumulative probabilities within a range of values can be computed on Casio graphing calculators.

Casio (9750/9850/9860)								
Computing probabilities using the Ncd function								
To compute P ($x \le 9,000$ hours) 1) From the menu screen select \square and press EXE 2) Press F5 (to choose "DIST") 3) Press F1 (to choose "NORM") 4) Press F2 (to choose "Ncd") 5) Enter values Lower bound = 0 Upper bound = 9,000 $\sigma = 1,000$ $\mu = 10,000$	 6) Arrow down to highlight "Execute" 7) Press F1 (to choose "CALC") P =0.68268949 z:Low=-1 z:UP =1 NOTE: For information on what "z:" means, see section on standard normal distribution and z-scores.							

After recording the probabilities for the Bulb Emporium as shown in Figure 3, Donna calculated the same interval probabilities for Lights-R-Us and recorded them.

Figure 3

	Bulb Emporium ($\sigma = 1000$)	Lights-R-Us ($\sigma = 750$)
$P(x \le 9,000 \text{ hours})$	0.1587	0.0912
$P(9,000 \le x \le 11,000 \text{ hours})$	0.6827	0.8176
$P(x \ge 11,000 \text{ hours})$	0.1587	0.0912

In analyzing her data, Donna noticed that a lightbulb at Bulb Emporium had a higher probability of lasting less than 9,000 hours than a bulb at Lights-R-Us. However, a lightbulb at Bulb Emporium also had a higher probability of lasting longer than 11,000 hours than a bulb at Lights-R-Us. She determined that these two statistical probabilities offset one another. The middle interval of data showed that a bulb at Bulb Emporium had a lower probability of lasting between 9,000 and 11,000 hours, inclusively, than a bulb at Lights-R-Us. She therefore chose to purchase lightbulbs from Lights-R-Us.

Standard normal curve

z-Scores

A z-score, also called a standard score, is a measure of position derived from the mean and standard deviation of the data set. The z-score is a measure of how many standard deviations an element falls above or below the mean of the data set. The z-score has a positive value if the element lies above the mean and a negative value if the element lies below the mean. A z-score associated with an element of a data set is calculated by subtracting the mean of the data set from the element and dividing the result by the standard deviation of the data set.

z-score $(z) = \frac{x - \mu}{\sigma}$, where x represents an element of the data set, μ represents the mean of the data set, and σ

represents the standard deviation of the data set.

The standard normal curve is a normal distribution that has a mean of 0 and a standard deviation of 1. It is used to model z-scores obtained from normally distributed data. Prior to calculator technologies that can determine probabilities associated with normal distributions, the table of Standard Normal Probabilities, commonly referred to as a "z-table" was used to determine normal distribution probabilities. See pages 11 and 12 for a copy of the table of Standard Normal Probabilities provided for the Algebra II End-of-Course (EOC) SOL test.

Given a z-score (z), the table of Standard Normal Probabilities shows the area under the curve to the left of z. This area represents the proportion of observations with a zscore less than the one specified. Table rows show the zscore's whole number and tenths place. Table columns show the hundredths place. In the table of Standard Normal Probabilities provided for the state EOC SOL test in Algebra



II, the cumulative probability from negative infinity to the z-score appears in table cells. Other tables of Standard Normal Probabilities show probabilities from the mean to the z-score.

Interpreting values from the table of Standard Normal Probabilities

A z-score associated with an element of a normal distribution is computed to be 1.23. The probability from the table of Standard Normal Probabilities associated with a z-score of 1.23 can be determined as indicated in Figure 4. The probability can be used differently based upon the context of the question.

- The probability that a data value will fall below the data value associated with a z-score of 1.23 is 0.8907 (89.07%).
- The data value associated with a z-score of 1.23 falls in the 89th percentile. This means that 89 percent of the data in the distribution fall below the value associated with a z-score of 1.23.
- The probability that a value from the data set will fall above this value is 1 0.8907 = 0.1093 (10.93%).







Figure 4 shows the cumulative probability associated with a z-score of 1.23 using the table of standard normal probabilities.

From Example 3, Donna could have determined $P(x \le 9,000 \text{ hours})$ for Lights-R-Us using the table of Standard Normal Probabilities.

Given $\mu = 10,000$ and $\sigma = 750$, the z-score for 9,000 hours can be computed using the formula for z-score.

$$z = \frac{9000 - 10000}{750} = -1.33.$$

In the table of Standard Normal Probabilities, z = -1.33 is associated with a cumulative probability of 0.0918, a value very close (difference of 0.0006) to the one found for Lights-R-Us in Figure 3, using a graphing calculator.

Example 4

The ACT[®] is an achievement test given nationally with normally distributed scores. Amy scored a 31 on the mathematics portion of her 2009 ACT[®]. The mean for the mathematics portion of the ACT[®] in 2009 was 21.0 and the standard deviation was 5.3. What percent of the population scored higher than Amy on the mathematics portion of the ACT[®]?

There are two "typical" approaches to find the percentage of the population that scored higher than Amy.

Solution A – Find the z-score and associated cumulative probability using the table of standard normal probabilities and subtract the cumulative probability from 1.

Calculate the z-score. $z = \frac{31-21.0}{5.3} = 1.89$. Look up the cumulative probability associated with a z-score of 1.89 on the table of standard normal probabilities. The probability of a test taker scoring a 31 (z-score = 1.89) is 0.9706 or 97 percent. This means that Amy scored in the 97th percentile and only 2.94 percent (1 – 0.9706 = 0.0294) scored higher than Amy.

Solution B – Find the probability of scoring higher than 31 using the graphing calculator.

The Texas Instruments (TI-83/84) graphing calculators can compute the percentage of the population scoring higher than 31 using the syntax normalcdf(31,43,21.0, 5.3), given lower bound = 31, upper bound = 43, $\mu = 21.0$, and $\sigma = 5.3$. The resulting value is 0.0296 or 2.96 percent (see page 5). Note: When choosing an upper bound (43 in this case), choose a number that will encompass all data values above the mean. The number 43 (rounded to 43 using $\mu + 4 \sigma = 21+21.2 = 42.2$) is far enough above the mean to be a "safe" upper bound.

The Casio (9750/9850/9860) graphing calculators can compute the percentage of the population that score above a 31 by using the operation "Ncd" or "Normal C.D." with a lower bound = 31, upper bound = 43, σ = 5.3, and μ = 21.0 (see page 5). The calculated percentage of the population scoring above a 31 = 0.0296 or 2.96 percent. The Casio 9750/9860 can "DRAW" the representative area under the curve in terms of z-scores by pressing **F6**.



Example 4 extension

Amy took the ACT[®] and scored 31 on the mathematics portion of the test. Her friend Stephanie scored a 720 on the mathematics portion of her 2009 SAT[®]. Both the SAT[®] and the ACT[®] are achievement tests given nationally with scores that are normally distributed. The mean for the mathematics portion of the SAT[®] in 2009 was 515 and the standard deviation was 116. For the ACT[®], the mean was 21 and the standard deviation was 5.3. Whose achievement was higher on the mathematics portion of their national achievement test?

Solution

Since these two national achievement tests have different scoring scales, they cannot be compared directly. One way to compare them would be to find the cumulative probability (percentile) of each score using the associated z-score.

Amy's z-score is 1.89 (97th percentile) and Stephanie's z-score is 1.77 (96th percentile). Therefore, Amy scored slightly higher than Stephanie on the mathematics portions of their respective national achievement tests.

Explorations with students

- Note: All exploration questions should be in the real-world context of normally distributed data sets.
- 1. Given a normally distributed data set with a specified mean and standard deviation, explain how the number of values expected to be above or below a certain value can be determined.
- 2. Given normally distributed data, explain how you can determine how many values or what percentage of values are expected to fall within one, two or three standard deviations of the mean.
- 3. Compare and contrast graphs of normal distributions that have the same mean but different standard deviations or different means and the same standard deviation.
- 4. Given a normally distributed data set with a specified mean and standard deviation, explain how to determine the probability and/or area under the curve for
 - an element that has a value greater than a given value;
 - an element that has a value less than a given value; or
 - an element that has a value between two given values.
- 5. Given the mean and standard deviation of two different normally distributed data sets, and a value from each data set, compare the values using their corresponding z-scores and percentiles.
- 6. Given normally distributed data with specified mean and standard deviation, determine the probability that a randomly selected value will have a z-score within a certain range of values.

Standard Normal Probabilities



Table entry for z is the area under the standard normal curve to the left of z.

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	.02 44	.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	. 1 251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	. 1 469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	. 1 711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	. 1 977	.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

Technical Assistance Document for AII.11

Standard Normal Probabilities



Table entry for z is the area under the standard normal curve to the left of z.

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

Technical Assistance Document for AII.11