

FORTY-SIXTH ANNUAL MATHEMATICS CONTEST
sponsored by
THE TENNESSEE MATHEMATICS TEACHERS' ASSOCIATION

Statistics 2002

Prepared by:

Steve Gadbois
Memphis, TN

Reviewed by:

Mathematics Faculty
Austin Peay State University
Clarksville, TN 37044

Scoring formula: $4R - W + 40$

DIRECTIONS:

Do not open this booklet until you are told to do so.

This is a test of your competence in high school mathematics. For each problem, determine the best answer and indicate your choice by making a heavy black mark in the proper place on the separate answer sheet provided. You must use a pencil with a soft head (No. 2 lead or softer).

This test has been constructed so that most of you are not expected to answer all of the questions. Do your best on the questions you feel you know how to work. You will be penalized for incorrect answers, so wild guesses are not advisable.

If you change your mind about an answer, be sure to erase completely. Do not mark more than one answer for any problem. Make no stray marks of any kind on the answer sheet. The answer sheets will not be returned to you. If you wish a record of your performance, mark your answers in this booklet also. You will keep the booklet after the test is completed.

When told to do so, open your test booklet and begin. You will have exactly 80 minutes to work.

Contributors to TMTA for the Annual Mathematics Contest:

Dr. Hal Ramer, President, Volunteer State Community College, Gallatin, Tennessee
Donnelley Printing Company, Gallatin, Tennessee
TRW Commercial Steering Division, Lebanon, Tennessee
Wright Industries, Inc., Nashville, Tennessee

1. In a stemplot, if 2.7436 is represented by $4|4$, which of the following could be represented by $4|2$?

- (a) 4.2
- (b) 42
- (c) 2.7426
- (d) 2.7416
- (e) 2.42

2. The reading comprehension level of 80 school age children was observed. Of the lowest 20 scores, 17 were of children who lived at or below the poverty level. Based on this, which one of the following five statements can be made?

- (a) 85% of children in poverty have poor reading skills.
- (b) Living in poverty causes lower reading skills.
- (c) Children of wealthy parents are naturally more intelligent.
- (d) There may be an association between income level and reading skill.
- (e) 25% of children read below grade level.

3. An experiment was designed to study the effects of a new fertilizer on the growth of a certain flower. Flowers of this species were placed by a random assignment in one of the following eight categories:

	no sun	minimal sun	moderate sun	full sun
fertilizer				
no fertilizer				

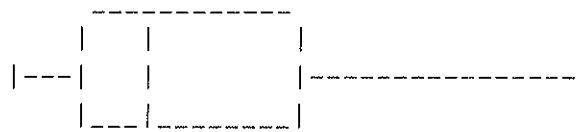
Which one of the following five statements is true of the design of this experiment?

- (a) It is completely randomized.
- (b) It is blocked by sunlight with four levels.
- (c) It is blocked by fertilizer.
- (d) It is blocked by fertilizer and sunlight.
- (e) It is not blocked at all.

4. One study shows that overall, males are more likely to obtain a college degree than females. But, when the data are broken down by socioeconomic status, in each group females are more likely to obtain a college degree than males. What term is used to describe this reversal of the direction of an association when lurking variables are taken into account?

- (a) negative association
- (b) Simpson's paradox
- (c) the placebo effect
- (d) blocking
- (e) Bayes theorem

5. Suppose that some data has the following boxplot.



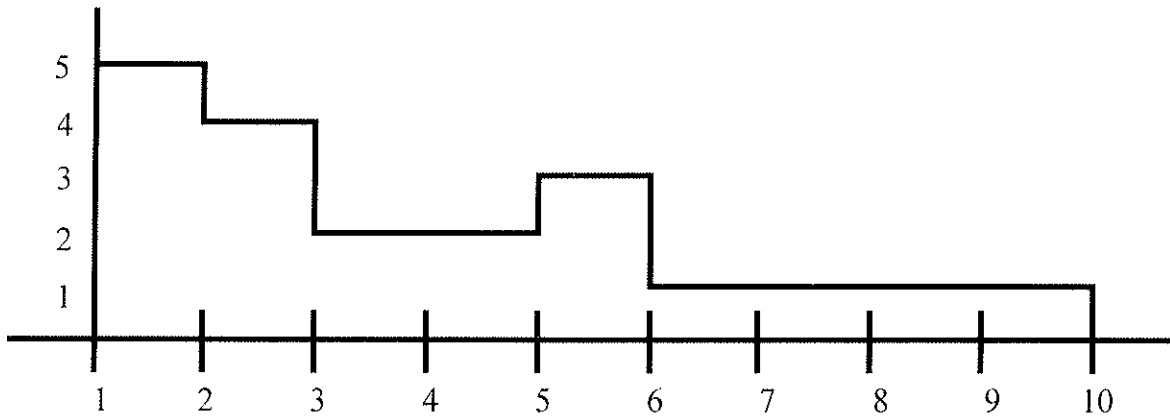
Consider the following three statements about the data.

- I. There are more values above the median than there are below it.
- II. The median is less than the mean.
- III. The mean is less than the median.

Which of the above statements are true?

- (a) I only
- (b) II only
- (c) III only
- (d) I and II only
- (e) I and III only

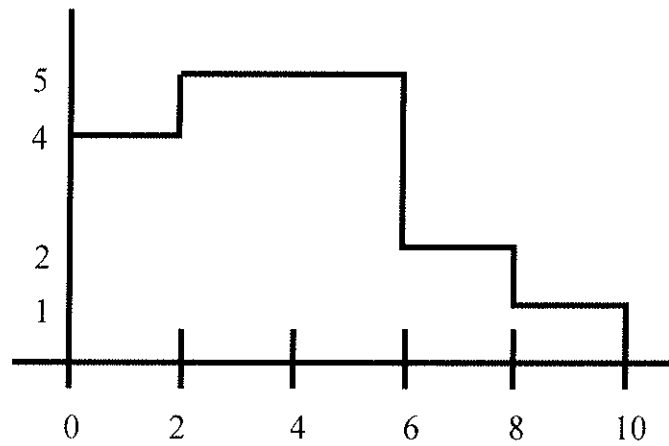
6. Consider the following histogram.



What is the approximate percentage of all values greater than 2 and less than 8?

- (a) 65%
- (b) 70%
- (c) 75%
- (d) 80%
- (e) 85%

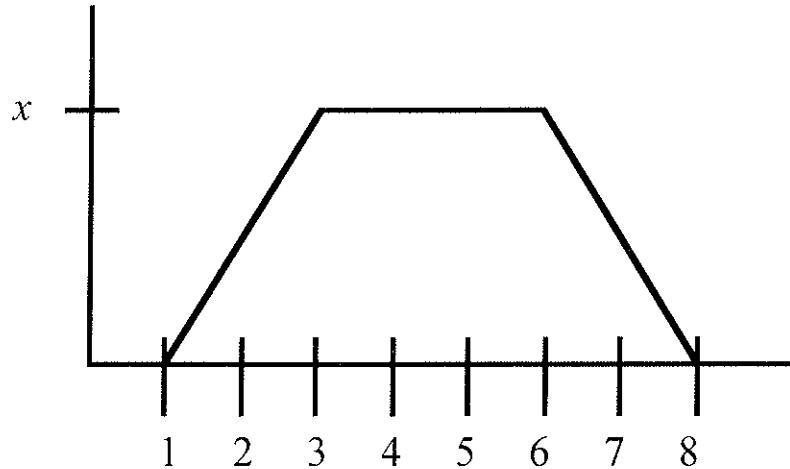
7. Consider the following histogram.



Which one of the five statements is the strongest true statement about \tilde{x} , the median of the data represented by the histogram?

- (a) $\hat{x} \leq 6$
- (b) $\tilde{x} \leq 5$
- (c) $\tilde{x} \leq 4$
- (d) $2 < \hat{x} \leq 4$
- (e) $4 < \tilde{x} \leq 6$

8. The following graph depicts a probability density function.



What is the value of the constant x ?

- (a) 1
- (b) $1/5$
- (c) $2/11$
- (d) $1/7$
- (e) $1/8$

9. Consider the following five general statements.

- I. When a poll is wrong, it is because many of those polled changed their opinion after the poll and before the vote.
- II. If a pollster is lucky enough to be able to include an entire population in her sample, the "margin of error" will be 0.
- III. The mean and the standard deviation of a set of numbers have the same "units" or "dimensions".
- IV. The sum of residuals is 0.
- V. For bivariate (paired) data, the correlation and the slope of the regression (least-squares) line have the same sign.

Which of the above statements must be true?

- (a) II, III, IV, and V
- (b) I, III, IV, and V
- (c) I, II, IV, and V
- (d) I, II, III, and V
- (e) I, II, III, and IV

10. Suppose that some data have a mean of 5 and a standard deviation of 3. Suppose that a new data set is created by replacing each original data item with the number gotten by multiplying the original data item by 2 and then adding 10. (For example, if 7 were in the original data set, then it would be replaced with 24, since $7 \cdot 2 + 10 = 24$.) What are the mean and standard deviation of the new data?

- (a) 5 and 3
- (b) 20 and 6
- (c) 20 and 16
- (d) 30 and 6
- (e) 30 and 26

11. Each of one hundred seniors in a simple random sample of all seniors reveal their GPA. Which one of the following five statements must be true?

- (a) The variance of this sample cannot be computed since the population size is not given.
- (b) The variance of this sample equals the variance of the population.
- (c) The variance of this sample is smaller than the variance of the population.
- (d) The variance of this sample is an estimate of the variance of the population but may differ from the variance of the population.
- (e) The variance of this sample is an estimate of the variance in the sampling distribution of the means of all possible samples.

12. Suppose that a set of three different numbers is converted to the corresponding set of z -scores (i.e., standardized scores). Which one of the following sets of three numbers could be the result?

- (a) $\{-10 \text{ feet}, 0 \text{ feet}, 10 \text{ feet}\}$
- (b) $\{-1, 0, 1\}$
- (c) $\{-10, 0, 10\}$
- (d) $\{1, 2, 3\}$
- (e) $\{-1, -2, 3\}$

13. Joe and Bob are amateur distance runners. They compete often and compare their times on the 10 km races they run. Historically, Joe averages 54.0 minutes with a standard deviation of 6.5 minutes, while Bob averages 55.5 minutes with a standard deviation of 4.5 minutes. Last week they ran against each other again; Joe finished with a time of 52.0 minutes and Bob finished with a time of 53.5 minutes. Relative to their usual performance, which runner ran better compared to his normal level, and why?

- (a) Bob, because a time of 53.5 minutes is difficult to improve upon.
- (b) Joe, because Joe's times vary more than Bob's times.
- (c) They performed the same because they are both exactly two minutes under their normal times.
- (d) Joe, because the probability of Joe being two minutes under his average is less than the probability of Bob being two minutes under his.
- (e) Bob, because the probability of Bob being two minutes under his average is less than the probability of Joe being two minutes under his.

14. Again consider amateur distance runners Joe and Bob. In the 10 km run, Joe averages 54.0 minutes with a standard deviation of 6.5 minutes, while Bob averages 55.5 minutes with a standard deviation of 4.5 minutes. Assume that each player's times are normally distributed, and that the two players' times are independent. Which one of the following gives the probability that there is a difference of more than 2 minutes in their times? (*You need not complete the probability calculation.*)

- (a) $P(z > (2 - 1.5)/5.5)$
- (b) $P(z < (2 - 1.5)/5.5)$
- (c) $P[(-2 - 1.5)/\sqrt{62.5} < z < (2 - 1.5)/\sqrt{62.5}]$
- (d) $1 - P[(-2 - 1.5)/\sqrt{62.5} < z < (2 - 1.5)/\sqrt{62.5}]$
- (e) $1 - P(-2/\sqrt{62.5} < z < 2/\sqrt{62.5})$

15. A certain college requires a score of 900 on the GBT test for admission, but it will also accept an equivalent grade on the MRST test. The mean score on the GBT is 1020 and the standard deviation is 140; the mean score on the MRST is 21 and the standard deviation is 4.7. What is the minimum score on the MRST that the college will accept?

- (a) 16
- (b) 17
- (c) 18
- (d) 19
- (e) 20

16. In a study comparing statistics grades with success in a particular job, a researcher found the coefficient of correlation to be 0.70. What percentage of the variation in success on the job can be explained by the variation in the statistics grades?

- (a) 9%
- (b) 30%
- (c) 49%
- (d) 70%
- (e) Not enough information is given.

17. Four data points, given in the form (x, y) , are $(1, 1)$, $(2, 3)$, $(3, 4)$, and $(4, 6)$. The equation of the regression (least-squares) line is $\hat{y} = -0.5 + 1.6x$. (*You need not verify this.*) What is the residual corresponding to $x = 4$?

- (a) 0.1
- (b) 1.1
- (c) 3.0
- (d) 5.9
- (e) 6.0

18. The least-squares regression line is fit to a set of data, and one of the data points has a negative residual. Which one of the following five statements must be true?

- (a) The data point must lie near the right edge of the scatterplot.
- (b) The data point must lie above the regression line.
- (c) The data point must lie below the regression line.
- (d) The correlation between the response and explanatory variables (i.e., the dependent and independent variables) must be negative.
- (e) The correlation between the response and explanatory variables (i.e., the dependent and independent variables) must be positive.

19. The owner of eight pizza shops did a study to determine the relationship between the number of students in an area and the number of pizzas he sold. The following table represents his results.

Shop number	1	2	3	4	5	6	7	8
Students (in 1000s)	5	3	22	8	11	3	17	9
Pizzas sold (in 100s)	14	6	31	13	18	10	20	14

Based on a simple linear regression, approximately how many pizzas could he expect to sell to a new area with 19,000 students?

- (a) 2500
- (b) 2550
- (c) 2600
- (d) 2650
- (e) 2700

20. Suppose that one simultaneously flips a fair coin once and tosses a fair six-sided die once. What is the probability of getting a “heads” *or* a “3”?

- (a) $1/12$
- (b) $1/6$
- (c) $1/2$
- (d) $7/12$
- (e) $2/3$

21. On this test of 40 questions with five possible answers for each, what is the probability of scoring exactly 8 correct if on every question, you have absolutely no idea and randomly guess?

- (a) .156
- (b) .173
- (c) .316
- (d) .437
- (e) .594

22. If you are told that at least one of three dogs is male, what is the probability that exactly two of the three dogs are female?

- (a) 0
- (b) $1/3$
- (c) $3/7$
- (d) $1/2$
- (e) 1

23. In this question, A^c represents the complement of the event A (sometimes also written as “not A ” or \bar{A}) and $A \cap B$ represents the intersection of the events A and B (sometimes also written as “ A and B ” or as AB).

Suppose that $P(A^c) = 0.2$, $P(B) = 0.4$, and $P(A \cap B) = 0.32$, and consider the following four statements.

I. $P(A) = 0.4$

II. $P(A) = 0.8$

III. Events A and B are independent.

IV. Events A and B are disjoint.

Which of the above statements are true?

- (a) I only
- (b) II only
- (c) I and IV only
- (d) II and III only
- (e) III and IV only

24. If two events are mutually exclusive (disjoint) and each has nonzero probability, then which one of the following statements is true of these two events?

- (a) The events are independent.
- (b) The events are not independent.
- (c) The events are neither dependent nor independent.
- (d) The events have probabilities that cannot be computed.
- (e) The events have probabilities that sum to a value greater than one.

25. A weather forecaster says there is a 30% chance of rain for each of the next three days. If you assume that the forecast is accurate and that the outcomes for each day are independent, what is the probability that there will be no rain at all during the next three days?

- (a) 0.027
- (b) 0.300
- (c) 0.343
- (d) 0.657
- (e) 0.700

26. In an episode of the original TV show “Star Trek”, some aliens are holding some members of the crew of the Starship Enterprise captive. One of these aliens states that the probability is 87% that Dr. McCoy will die, while the probability is 93% that Mr. Spock will suffer severe brain damage. Assuming that these figures are accurate and that the two events (Dr. McCoy dying and Mr. Spock suffer severe brain damage) are independent, what is the probability that *neither* event will occur?

- (a) 0.0091
- (b) 0.07
- (c) 0.13
- (d) 0.2
- (e) 0.9909

27. Suppose that X and Y are independent random variables with probability distributions given by the following two tables.

X	=	1	2	3
$P(X)$	=	0.2	a	0.4

and

Y	=	1	2	3
$P(Y)$	=	b	0.3	c

(Here, a , b , and c are constants.) Given that $P(X = 2, Y = 1) = 0.1$, what is the value of c ?

- (a) 0.04
- (b) 0.25
- (c) 0.40
- (d) 0.45
- (e) The value of c cannot be determined from the given information.

28. In the carnival game Chuck-a-Luck, the player chooses one of the numbers 1, 2, 3, 4, 5, 6. Then three ordinary (and presumably fair) dice are rolled. If the player's number does not appear on any of the dice, the player loses \$1 (i.e., wins $-\$1$). If the player's number appears on one, two, or three of the dice, the player wins \$1, \$2, or \$3 respectively. The probabilities of having the number appear on one, two, or three of the dice are $75/216$, $15/216$, or $1/216$ respectively. (*You need not verify these probabilities.*) What is the player's approximate expected (average) result per game?

- (a) 0 cents
- (b) -7.87 cents
- (c) 42.13 cents
- (d) -50 cents
- (e) 57.87 cents

29. The dean of a small southwestern college determined that the size of lecture classes over the last five years could be approximated by a uniform distribution of between 41 and 90 students (inclusive). What is the approximate probability of a class having 80 students or less?

- (a) 0.02
- (b) 0.10
- (c) 0.20
- (d) 0.80
- (e) 0.90

30. If the heights of adult men are approximately normally distributed about a mean of 70 inches with a standard deviation of 3 inches, what is the approximate height of a man at the 98th percentile?

- (a) 76 inches
- (b) 77 inches
- (c) 78 inches
- (d) 79 inches
- (e) 80 inches

31. Assume that the weights of an individual potato of a certain type are normally distributed with a mean of 6 ounces and a standard deviation of 1 ounce and that the weights of individual potatoes are independently distributed. What would be the distribution of the total weight of a bag of 100 potatoes?

- (a) normal with mean 6 ounces and standard deviation 0.1 ounce
- (b) normal with mean 6 ounces and standard deviation 1 ounce
- (c) normal with mean 60 ounces and standard deviation 1 ounce
- (d) normal with mean 600 ounces and standard deviation 100 ounces
- (e) normal with mean 600 ounces and standard deviation 10 ounces

32. Let

$$i = P(-1 < z < 1),$$

$$j = P(-2 < z < 0), \text{ and}$$

$$k = P(0 < z < 3).$$

From smallest to largest, what is the proper order of the three quantities i , j , and k ?

- (a) $i < j < k$
- (b) $j < i < k$
- (c) $j < k < i$
- (d) $k < i < j$
- (e) $k < j < i$

33. Suppose that a random sample of 100 independent measurements is taken from a population having mean $\mu = 5$ and standard deviation $\sigma = 7$. What is the approximate probability that the sample mean \bar{x} will be between 4.0 and 5.5?

- (a) 0.03
- (b) 0.07
- (c) 0.09
- (d) 0.16
- (e) 0.69

34. Consider the following three general statements about the width of confidence intervals. All else being equal.

I. a confidence interval for a sample with smaller size is wider than a confidence interval for a sample with larger size.

II. a confidence interval for a sample with smaller standard deviation is wider than a confidence interval for a sample with larger standard deviation.

III. a confidence interval for lower confidence (e.g., 90%) is wider than a confidence interval for higher confidence (e.g., 99%).

Which of the above statements must be true?

- (a) I only
- (b) II only
- (c) III only
- (d) I and III only
- (e) II and III only

35. A statistics professor routinely recycles his tests every four years. His records showed that grades on his final exam four years ago were better than this year, and he wondered if the performance of his students had changed since four years ago. He collected the following results.

	Four Years Ago	This Year
Mean	88.0	82.0
Variance	112.5	54.0
Sample Size	45	36

What is the 95% confidence interval for the difference between the two population means?

- (a) $(-9.92, -2.08)$, i.e., from -9.92 to -2.08
- (b) $(-8.44, -3.56)$
- (c) $(-3.92, 3.92)$
- (d) $(-13.84, 1.84)$
- (e) $(-24.23, 12.23)$

36. If a confidence interval for a proportion is computed to be $(0.54, 0.62)$ and the sample size is $n = 2000$, what is x , the number of "successes" in the sample?

- (a) 0.58
- (b) 58
- (c) 1160
- (d) It cannot be determined because the confidence level is not given.
- (e) It cannot be determined because the critical value is not given.

37. When making the case for a new TMTA Statistics contest, proponents claimed that the proportion of Tennessee high schools offering an appropriate course was at least 0.20, while opponents felt that the proportion was less. If p represents the true proportion of such schools offering such a course, which of the following best give the null hypothesis (H_0) and alternative hypothesis (H_a) that opponents could test?

- (a) $H_0 : p \leq 0.20, H_a : p = 0.20$
- (b) $H_0 : p > 0.20, H_a : p = 0.20$
- (c) $H_0 : p < 0.20, H_a : p \geq 0.20$
- (d) $H_0 : p = 0.20, H_a : p > 0.20$
- (e) $H_0 : p \geq 0.20, H_a : p < 0.20$

38. The average life expectancy of a line of electric shavers is 9.3 years. Management believes that an improvement in the production process will increase the live expectancy. If μ represents the true average life expectancy with the improvement, which of the following best give the null hypothesis (H_0) and alternative hypothesis (H_a) that management could use to test the validity of management's belief?

- (a) $H_0 : \mu < 9.3; H_a : \mu \geq 9.3$
- (b) $H_0 : \mu \leq 9.3; H_a : \mu > 9.3$
- (c) $H_0 : \mu > 9.3; H_a : \mu \leq 9.3$
- (d) $H_0 : \mu \geq 9.3; H_a : \mu < 9.3$
- (e) $H_0 : \mu = 9.3; H_a : \mu \neq 9.3$

39. In a hypothesis test, what is by definition the name of the probability of rejecting a false null hypothesis when an alternate hypothesis is true?

- (a) p -value
- (b) α (i.e., alpha)
- (c) β (i.e., beta)
- (d) $1 - \beta$
- (e) confidence level

40. A researcher tested the price of unleaded gas to determine if average prices varied from state to state. She recorded the price at the same number of stations in several states and then obtained the following analysis of variance table.

Source	SS	df	MS	F
Between	106	4	26.5	4.14
Within	96	15	6.4	
Total	202	19		

Also, a table of critical values for the F distribution provides the value $F_{0.05} = 3.06$. Which one of the following *cannot be concluded* from this?

- (a) Stations were tested in five different states.
- (b) Four stations were tested in each state.
- (c) At least two states had the same average price.
- (d) The null hypothesis was that the average price in all states was the same.
- (e) The null hypothesis should be rejected at the $\alpha = 0.05$ level.

Forty-sixth Annual Mathematics Contest
Tennessee Mathematics Teachers Association
STATISTICS 2002

Notation:

$P(A)$ represents the probability of the event A

z always represents a quantity having a standard normal (i.e., Gaussian) distribution

Some possibly useful formulas:

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Only two statistical tables are needed and provided for this contest exam: the "standard normal" table of probabilities and the t -table of critical values. The form of the standard normal table provided has probabilities of the form $P(0 < z < c)$, where c is a constant ranging from 0 to 3.49. *Each problem on this contest exam has an ordinary solution not requiring any other statistical tables.*

CRITICAL VALUES OF "STUDENT'S T" DISTRIBUTION

Critical values t_p satisfy $p = P(t \geq t_p)$.

d.f.	$t_{.250}$	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.0025}$	$t_{.001}$	$t_{.0005}$
1	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.598
3	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
inf.	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

