

STATISTICS 100 EXAM 2**Spring 2016****PRINT** NAME _____

(Last name)

(First name)

***NETID** _____**CIRCLE SECTION:** **L1 (Laska MWF)** **L2 (Tues/Thurs)** **Robin Tu**Write answers in appropriate blanks. When no blanks are provided **CIRCLE** your answers.**SHOW WORK** when requested, otherwise no credit.

Do NOT use scrap paper.

Make sure you have all 7 pages including the normal table (14 problems).**DO NOT WRITE BELOW THIS LINE**

The numbers written in each blank below indicate how many points you missed on each page. The numbers printed to the right of each blank indicate how many points each page is worth.

Page 1 _____ 27

Page 2 _____ 17

Page 3 _____ 12

Page 4 _____ 18

Page 5 _____ 12

Page 6 _____ 14

Extra Credit _____

Total Score _____***Karle Laska's Sections: There is NO class Thursday and Friday! Have a great weekend!****Scores will be posted in Compass early Friday morning ☺**

Question 1: Follow the steps below and draw a boxplot for the data given. (10 points total)

Here is the data from 10 random students this semester who answered the survey question: "How many people have you been in a serious relationship with?": 7, 8, 0, 5, 10, 15, 6, 9, 6, 9

Step 1: Find the median, Q1, Q3, and the IQR.
Circle your answers.

(1 point) Median=

(1 point) Q1=

(1 point) Q3=

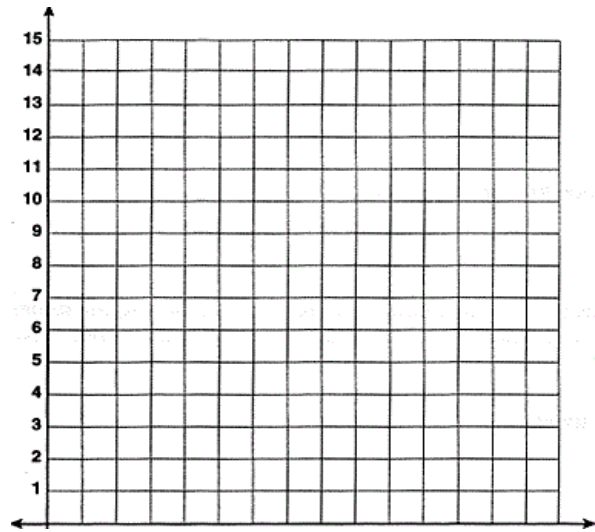
(1 point) IQR=

Step 2: Check for outliers. Fill in the blanks with yes or no.

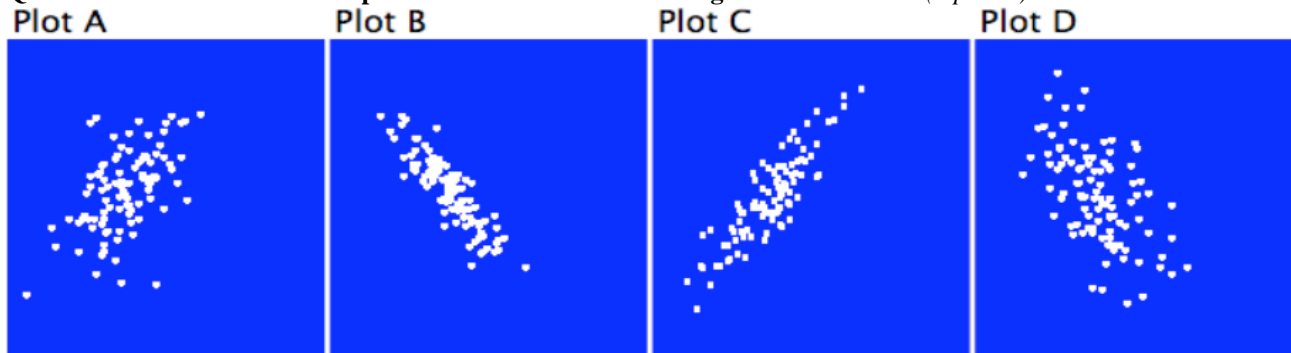
(1 point) Are there low outliers? _____
Yes/No

(1 point) Are there high outliers? _____
Yes/No

Step 3: Draw the boxplot on the axis below. (4 points)



Questions 2: Match the scatter plots below to the correlations given in the table. (4 points)



Write the letter of the plot next to the correlation coefficient that is closest to.

i)r=0.88 _____ ii)r=0.57 _____ iii)r= -0.88 _____ iv)r= -0.54 _____

Question 3: Part a) Compute the correlation coefficient (r) by filling in the table below including the totals.

The average of X=4 and the SD of X=2. The average of Y=4 and the SD of Y=2.

X	Y	Z-score for X	Z-score for Y	Products	Plot the 5 points below:
1	4				
3	7				
4	3				
5	1				
7	5				
Totals		Total should = _____	Total should = _____	Total = _____	

Part b) (1 point) The correlation coefficient $r =$ _____ (Fill in the blank and round to 2 decimal places)

Question 4 (5 points total)

For each of the following pairs of variables, check the box that best describes its correlation coefficient r .

Check only one box per row! READ THIS- Each column will be used exactly once!

	Exactly +1	Between 0 and +1	About 0	Between 0 and -1	Exactly -1
X is always 3 times Y.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Number of pets you have and Height	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Number of hours spent studying and Exam 2 score	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
X and Y always add up to 30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
How many calories you burn while exercising and how much you weigh	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Question 5 (2 points)

Suppose I calculate the average on Exam 1 and the average on Exam 2 for each of our 3 Stat 100 sections and find that the correlation between the 3 pairs of averages is 0.95. Can I conclude that if I drew a scatter plot of all 900 individual students' exam 1 and exam 2 scores, that the correlation for the 900 pairs would also be 0.95? (2 pts.)

Choose one:

- i) Yes, since the class averages are computed from the 900 students' scores, it follows that the correlation for the 900 pairs would also be 0.95.
- ii) No, the correlation for the 900 pairs of averages is likely to be **higher** than 0.95 since including all the information from the individuals within each class would increase accuracy and reduce scatter.
- iii) No, the correlation for the 900 pairs is likely to be **lower** than 0.95, since the scatter plot of the 3 pairs only shows the scatter between the 3 sections while the scatter plot of the 900 pairs would also show all the scatter within each section.

Question 6 X and Y are 2 sets of numbers with a correlation coefficient of $r = 0.8$. (4 points total)

Fill in the 4 blanks below with **numbers** (NOT with phrases like "increase" or "stay the same".)

- a. If 1 is subtracted from all the X values the new correlation coefficient would be= _____
- b. If all the original X values are multiplied by **negative** 0.5, the new correlation coefficient would be = _____.
- c. If all the original X values are multiplied by 0.5, the new correlation coefficient would be = _____.
- d. If all the original X and Y values are converted to z scores the new correlation coefficient would be = _____.

Question 7: The table below gives the 5 summary statistics of the 724 students who responded to our survey questions: "What is your ACT score?" and "What is your GPA in college so far?" (8 points total)

	Average	SD
ACT	27.8	3.5
GPA	3.3	0.4

Correlation Coefficient (r) = 0.3

a) Make regression estimates for the two students below by filling in the blanks in the table. For Hillary, you're given ACT scores and asked to predict GPA. For Bernie, you're given GPA and asked to predict ACT. Show work converting values to z scores and z scores to values as indicated. Don't round any of your answers!

ACT	ACT z-score	r	GPA z-score	GPA
Hillary ACT=32	$Z = \underline{\hspace{2cm}}$ (1 point)	$r = 0.3$	$Z = \underline{\hspace{2cm}}$ (1 point)	GPA= <u> </u> (1 point) Show work for full credit
ACT= <u> </u> (1 point) Show work for full credit	$Z = \underline{\hspace{2cm}}$ (1 point)	$r = 0.3$	$Z = \underline{\hspace{2cm}}$ (1 point)	Bernie GPA= 4.0

Given the following summary statistics from the previous page and the fact that $r=0.3$, answer parts b and c.

	Average	SD
ACT	27.8	3.5
GPA	3.3	0.4

b) What is the SD of the prediction errors when predicting GPA from ACT?

- i) 0 ii) .4 iii) $\sqrt{1-0.3^2} * 0.4$ iv) $\sqrt{1-0.3^2} * 3.5$ v) $0.3*(0.4/3.5)$ vi) $0.3*(3.5/0.4)$

c) What is the slope of the regression equation when predicting GPA from ACT?

- i) 0 ii) .4 iii) $\sqrt{1-0.3^2} * 0.4$ iv) $\sqrt{1-0.3^2} * 3.5$ v) $0.3*(0.4/3.5)$ vi) $0.3*(3.5/0.4)$

Question 8 (5 points total)

This question pertains to our new cheating detection system (See the cameras at the front of the room?). If you're cheating, there's a 99% chance of being detected. If you're not cheating, the system will incorrectly detect that you're cheating 6% of the time. Suppose 5% of all students in stat 100 actually cheat. Fill in the chart below for 2000 students and leave the answers for parts a and b as fractions.

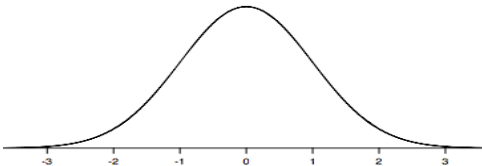
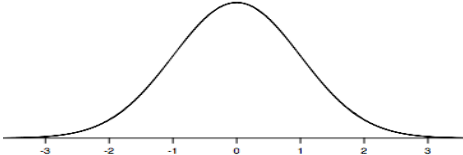
	Detected	Not Detected	Total
Cheating	(1 point)		(1 point)
Not Cheating	(1 point)		
Total			2000

a) (1 point) Given that a person is detected, what's the probability that they were actually cheating? _____

b) (1 point) Suppose a person is not cheating, what's the chance that they will be incorrectly detected? _____

Question 9 (10 points total)

a) If a husband is in the 10th percentile in weight where $r = 0.8$, what percentile would you estimate for his wife's weight? Solve by filling in the table below. (You may round areas and z-scores to fit the nearest line on the table.)

Husband's Weight Percentile	Husband Z	R=0.8	Wife Z	Wife's Weight Percentile
10th Percentile What middle area on the table should you look up to find the Z score? _____% (1 point)  Correctly mark the z-score and shade the area corresponding to the 10th percentile. (1/2 point for shading correctly)	Z = _____ (1 point)	r = 0.8	Z = _____ (1 point) Round your z-score to the nearest tenth.	Wife's Weight Percentile = _____ (1 point) Mark the Z score on the graph below. Shade the correct percentile.  Round the middle area given in the Normal Table to the nearest WHOLE number and then calculate the percentile. (1/2 point for shading correctly)

Suppose husbands' and wives' weights follow the normal curve but have different correlations in different countries.

b) Consider 5 countries where the correlation coefficients between the weights of husbands and wives are as given in the table below. If a husband is in the 25th percentile in weight, estimate his wife's percentile in weight for each country.

Percentile in Spatial Skills	r	Percentile in Social Skills
25 th	0.50	Choose One: 15 th 25 th 37 th 50 th 63 rd 75 th 85 th
25 th	1	Choose One: 15 th 25 th 37 th 50 th 63 rd 75 th 85 th
25 th	-0.50	Choose One: 15 th 25 th 37 th 50 th 63 rd 75 th 85 th
25 th	-1	Choose One: 15 th 25 th 37 th 50 th 63 rd 75 th 85 th
25 th	0	Choose One: 15 th 25 th 37 th 50 th 63 rd 75 th 85 th

Question 10 (13 points total)

The table below gives the 5 summary statistics from our survey data on the heights and weights of Stat 100 men. The scatter plot was roughly football shaped. Below are the 5 summary statistics.

	Average	SD
Height (in inches)	70	3
Weight (in pounds)	165	30

Correlation Coefficient (r)= **0.5**

- a) Find the regression equation for **predicting Weight** from height is: $Y = ______ X + ______$
- i) (2 points) First calculate the **slope**. Show work below. **Round to 2 decimal places!**
- ii) (2 points) Next, calculate the **y-intercept**. Show work below. **Round to 2 decimal places!**
- b) (2 points) Now use the equation you got above to estimate the weight of a guy in our class, Donald, who is 73". Show work, circle answer.
- c) (2 points) Suppose Donald actually weighs 153 pounds. What is his *residual or prediction error*? Show work below:
- d) (1 point) The average of the residuals (prediction errors) is _____. Fill in the blank with a number.
- e) (2 points) The regression equation for estimating height of the men class from weight is: Height = 0.05 * (Weight) + 61.75"
- i) Use the given regression equation to estimate the height of a guy in our class who weighs 195 lbs. (Show work. Circle answer. Don't round)
- ii) There's about a 68% chance that your estimate in (i) above is right to within _____ inches. Circle one:

$\sqrt{1 - 0.5^2} * 30$

$\sqrt{1 - 0.5^2} * 3$

0.5

30

3

- f) (2 points) Say another person, Marco, is 1.5 SD's below average in both height and weight. Which line does he fall on?

Choose one:

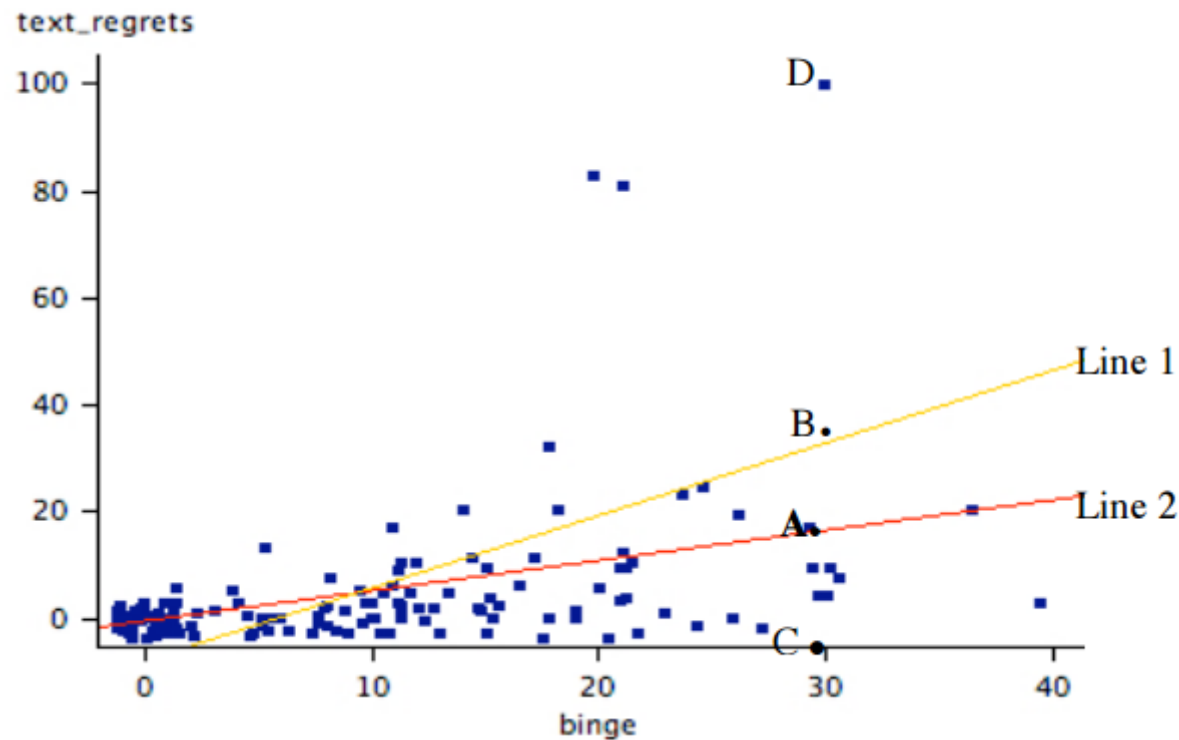
i) SD line

ii) Regression Line

iii) Both

iv) Neither

Question 11 (10 points total) The scatter plot below shows the responses of 156 students to 2 survey questions last spring: “What is the most you’ve ever drunk in one 24 hour period?” (Labeled “binge”) on the X axis and “How many text messages have you sent this semester that you regret sending?” (Labeled “text regrets”) on the Y axis.



- a) Look at Line 1 and Line 2 on the scatter plot above. One is the SD line and one is the regression line. Which is the regression line? Choose one: i) Line 1 ii) Line 2
- b) The correlation (r) between bingeing and text regrets is closest to
Choose one: i) -0.85 ii) 0.85 iii) -0.4 iv) 0.4
- c) The point of averages lies on... Choose one: i) Only Line 1 ii) Only Line 2 iii) Both Lines iv) Neither Line
- d) Look at students A, B, C and D on the graph. A is on Line 2, B is on Line 1, C is below both lines and D is above both. Their prediction errors (in no particular order) are: 0, 85, 15, and -15. Which student has which prediction error?
i) Student A has a prediction error of ... **Choose one:** a) 0 b) 85 c) 15 d) -15
ii) Student B has a prediction error of ... **Choose one:** a) 0 b) 85 c) 15 d) -15
iii) Student C has a prediction error of ... **Choose one:** a) 0 b) 85 c) 15 d) -15
iv) Student D has a prediction error of ... **Choose one:** a) 0 b) 85 c) 15 d) -15

Question 12 pertains to the table below which shows the responses of the 500 females in our class to the 2 survey questions: Do you believe sex before marriage is immoral?” and “What is your ethnicity?” (7 points total)

	White	Asian	African American	Latino	Mixed/Other	Totals
Yes, immoral	38	37	11	7	2	95
No, not immoral	260	27	37	33	26	383
Depends on Type of Sex	19	1	1	1	0	22
Totals	317	65	49	41	28	500

Suppose you randomly draw from these students:

- a) What is the chance of getting someone who is Latino?
i) 22/500 ii) 41/500 iii) 62/500 iv) 63/500 v) 117/500 vi) 37/95 vii) 37/65
- b) What is the chance of getting someone who answered “Depends on Type of Sex”?
i) 22/500 ii) 41/500 iii) 62/500 iv) 63/500 v) 117/500 vi) 37/95 vii) 37/65

	White	Asian	African American	Latino	Mixed/Other	Totals
Yes, immoral	38	37	11	7	2	95
No, not immoral	260	27	37	33	26	383
Depends on Type of Sex	19	1	1	1	0	22
Totals	317	65	49	41	28	500

c) What is the chance of drawing once & getting either a Latino or someone who answered “Depends on Type of Sex”?
 i) 22/500 ii) 41/500 iii) 62/500 iv) 63/500 v) 117/500 vi) 37/95 vii) 37/65

d) What is the chance you’ll get an Asian if you draw only from those who answered “Yes”?
 i) 102/500 ii) 62/500 iii) 63/500 iv) 117/500 v) 37/95 vi) 37/65

e) Draw 3 students with replacement. What is the chance that at least one of them answered “Yes”?
 i) $(95/500)^3$ ii) $(405/500)^3$ iii) $1-(95/500)^3$ iv) $1-(405/500)^3$ v) $(22/500)$ 3 vi) $95/500 + 95/500 + 95/500$

f) Draw 3 students without replacement. What is the chance that all 3 students are African American?
 i) $(49/500)^3$ ii) $49/500 * 48/499 * 47/498$ iii) $317/500 * 65/500 * 41/500$ v) $1-(49/500)^3$ vi) $317/500 * 65/499 * 41/498$

g) Draw 3 students without replacement. What’s the chance the first is African American, the second is Asian, and the third is Latino?
 i) $(49/500)^3$ ii) $49/500 * 48/499 * 47/498$ iii) $49/500 * 65/500 * 41/500$ v) $49/500 * 65/499 * 41/498$ vi) $1-(49/500)^3$

Question 13 pertains to a well-shuffled deck of 52 cards. A deck of cards has 4 suits: clubs, diamonds, hearts and spades. There are 13 cards in each suit: 2 through 10, jack, queen, king, ace. (3 points total)

a) Draw 2 cards **without replacement**. What is the chance that both cards are spades?
 i) 8/52 ii) $13/52 * 12/51$ iii) 16/52 iv) $13/52 * 13/52$ v) $4/52 * 4/52$ vi) $4/52 * 3/51$ vii) 17/52

b) Draw **one** card at random. What is the chance that it’s **either** a 4 or a Queen?
 i) 8/52 ii) 26/52 iii) 16/52 iv) $13/52 * 13/52$ v) $4/52 * 4/52$ vi) $4/52 * 3/51$ vii) 17/52

c) Draw **one** card at random. What’s the chance that it’s **either** a Queen or a Diamond?
 i) 8/52 ii) $13/52 * 12/51$ iii) 16/52 iv) $13/52 * 13/52$ v) $4/52 * 4/52$ vi) $4/52 * 3/51$ vii) 17/52

Question 14 pertains to rolling fair dice. (6 points total)

a) Two dice are rolled. What is the chance that the sum of the spots is 8?
 i) 2/36 ii) 3/36 iii) 4/36 iv) 5/36 v) $1/6 * 1/6$ vi) 7/36

b) One die is rolled 3 times. What is the chance of getting no 4’s?
 i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1-(5/6)^3$ iv) $1-(1/6)^3$ v) 3/6

c) One die is rolled 3 times. What is the chance of getting at least one 4?
 i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1-(5/6)^3$ iv) $1-(1/6)^3$ v) 3/6

d) What is the chance of rolling a die once and getting either a 2 or a 3?
 i) $(5/6)^2$ ii) $1/6 * 1/6$ iii) $1-(5/6)^2$ iv) $1/6 + 1/6 - 1/36$ v) $1/6 + 1/6$

e) What is the chance of rolling a die twice and getting a 2 on the first roll and a 3 on the second roll?
 i) $(5/6)^2$ ii) $1/6 * 1/6$ iii) $1-(5/6)^2$ iv) $1/6 + 1/6 - 1/36$ v) $1/6 + 1/6$

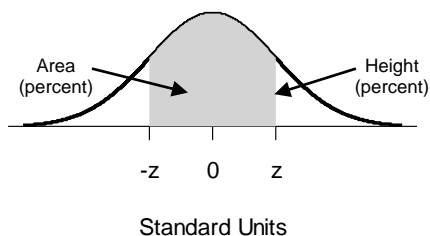
f) What is the chance of rolling a die twice and getting either a 2 on the first roll or a 3 on the second roll?
 i) $(5/6)^2$ ii) $1/6 * 1/6$ iii) $1-(5/6)^2$ iv) $1/6 + 1/6 - 1/36$ v) $1/6 + 1/6$

Extra Credit: You get 1 point for answering the first two questions correctly (all or nothing) and 1 point for answering the 3rd question correctly.

1. When is Karle Laska’s birthday? (month and day) _____
 2. Who did UIUC hire as the new football coach this week? (First and last name) _____

3. iPhone passwords are a combination of 4 numbers. After the 6th try, the phone locks itself for 1 minute. Knowing this, calculate the following probability: Suppose the smudge marks on your friend’s iPhone give away the four numbers, but you don’t know the order of those numbers.
 What is the probability that you unlock the phone in a single attempt? _____

STANDARD NORMAL TABLE



<i>z</i>	<i>Height</i>	<i>Area</i>		<i>z</i>	<i>Height</i>	<i>Area</i>		<i>z</i>	<i>Height</i>	<i>Area</i>
0.00	39.89	0.00		1.50	12.95	86.64		3.00	0.443	99.730
0.05	39.84	3.99		1.55	12.00	87.89		3.05	0.381	99.771
0.10	39.70	7.97		1.60	11.09	89.04		3.10	0.327	99.806
0.15	39.45	11.92		1.65	10.23	90.11		3.15	0.279	99.837
0.20	39.10	15.85		1.70	9.40	91.09		3.20	0.238	99.863
0.25	38.67	19.74		1.75	8.63	91.99		3.25	0.203	99.885
0.30	38.14	23.58		1.80	7.90	92.81		3.30	0.172	99.903
0.35	37.52	27.37		1.85	7.21	93.57		3.35	0.146	99.919
0.40	36.83	31.08		1.90	6.56	94.26		3.40	0.123	99.933
0.45	36.05	34.73		1.95	5.96	94.88		3.45	0.104	99.944
0.50	35.21	38.29		2.00	5.40	95.45		3.50	0.087	99.953
0.55	34.29	41.77		2.05	4.88	95.96		3.55	0.073	99.961
0.60	33.32	45.15		2.10	4.40	96.43		3.60	0.061	99.968
0.65	32.30	48.43		2.15	3.96	96.84		3.65	0.051	99.974
0.70	31.23	51.61		2.20	3.55	97.22		3.70	0.042	99.978
0.75	30.11	54.67		2.25	3.17	97.56		3.75	0.035	99.982
0.80	28.97	57.63		2.30	2.83	97.86		3.80	0.029	99.986
0.85	27.80	60.47		2.35	2.52	98.12		3.85	0.024	99.988
0.90	26.61	63.19		2.40	2.24	98.36		3.90	0.020	99.990
0.95	25.41	65.79		2.45	1.98	98.57		3.95	0.016	99.992
1.00	24.20	68.27		2.50	1.75	98.76		4.00	0.013	99.9937
1.05	22.99	70.63		2.55	1.54	98.92		4.05	0.011	99.9949
1.10	21.79	72.87		2.60	1.36	99.07		4.10	0.009	99.9959
1.15	20.59	74.99		2.65	1.19	99.20		4.15	0.007	99.9967
1.20	19.42	76.99		2.70	1.04	99.31		4.20	0.006	99.9973
1.25	18.26	78.87		2.75	0.91	99.40		4.25	0.005	99.9979
1.30	17.14	80.64		2.80	0.79	99.49		4.30	0.004	99.9983
1.35	16.04	82.30		2.85	0.69	99.56		4.35	0.003	99.9986
1.40	14.97	83.85		2.90	0.60	99.63		4.40	0.002	99.9989
1.45	13.94	85.29		2.95	0.51	99.68		4.45	0.002	99.9991

