Old Exam Questions for Practice
Econ 29000
Kevin R Foster, CCNY

1. Calculate the probability in the following areas under the Normal pdf with mean and standard deviation as given. You might usefully draw pictures as well as making the calculations. For the calculations you can use either a computer or a table.
    a. What is the probability, if the true distribution has mean -15 and standard deviation of 9.7, of seeing a deviation as large (in absolute value) as -1?
    b. What is the probability, if the true distribution has mean 0.35 and standard deviation of 0.16, of seeing a deviation as large (in absolute value) as 0.51?
    c. What is the probability, if the true distribution has mean -0.1 and standard deviation of 0.04, of seeing a deviation as large (in absolute value) as -0.16?
2. Using data from the NHIS, the National Health Interview Survey, we find the fraction of children who are female, who are Hispanic, and who are African-American, for two separate groups: those with and those without health insurance. Compute tests of whether the differences in the means are significant; explain what the tests tell us. *(Note that the numbers in parentheses are the standard deviations.)*

	with health insurance	without health insurance
female	0.4905	0.4811
	(0.49994) N=7865	(0.49990) N=950
Hispanic	0.2587	0.5411
	(0.43797) N=7865	(0.49857) N=950
African American	0.1785	0.1516
	(0.38297) N=7865	(0.35880) N=950

3. For the ATUS dataset, use "Analyze \ Descriptive Statistics \ Crosstabs" to create a joint probability table showing the fractions of males/females about the amount of time spent on the computer vs watching TV (if either or both are above average). Find and interpret the joint probabilities and marginal probabilities. Do this for age groups as well.
4. For the NHIS dataset, use "Analyze \ Descriptive Statistics \ Crosstabs" to create a joint probability table showing the fractions of males/females with or without health insurance. Find and interpret the joint probabilities and marginal probabilities. Do this for age groups and race/ethnicity as well.
5. For the NHANES dataset, use "Analyze \ Descriptive Statistics \ Crosstabs" to create a joint probability table showing the fractions of males/females smoking, drinking alcohol, doing drugs, and having lots of sex (define each term as you see fit; explain and justify). Find and interpret the joint probabilities and marginal probabilities. Do this for age groups and race/ethnicity as well.
6. You are in charge of polling for a political campaign. You have commissioned a poll of 300 likely voters. Since voters are divided into three distinct geographical groups (A, B and C), the poll is subdivided into three groups with 100 people each. The poll results are as follows:

	total	A	B	C
number in favor of candidate	170	58	57	55
number total	300	100	100	100

    Note that the standard deviation of the sample (not the standard error of the average) is given.
    a. Calculate a t-statistic, p-value, and a confidence interval for the main poll (with all of the people) and for each of the sub-groups.
    b. In simple language (less than 150 words), explain what the poll means and how much confidence the campaign can put in the numbers.
    c. Again in simple language (less than 150 words), answer the opposing candidate's complaint, "The biased media confidently says that I'll lose even though they admit that they can't be sure about any of the subgroups! That's neither fair nor accurate!"
7. Calculate the probability in the following areas under the Standard Normal pdf with mean of zero and standard deviation of one. You might usefully draw pictures as well as making the calculations. For the calculations you can use either a computer or a table.
    a. What is the probability, if the true distribution is a Standard Normal, of seeing a deviation from zero as large (in absolute value) as 1.9?
    b. What is the probability, if the true distribution is a Standard Normal, of seeing a deviation from zero as large (in absolute value) as -1.5?
    c. What is the probability, if the true distribution is a Standard Normal, of seeing a deviation as large (in absolute value) as 1.2?

8. Calculate the probability in the following areas under the Normal pdf with mean and standard deviation as given. You might usefully draw pictures as well as making the calculations. For the calculations you can use either a computer or a table.
    a. What is the probability, if the true distribution has mean -1 and standard deviation of 1.5, of seeing a deviation as large (in absolute value) as 2?
    b. What is the probability, if the true distribution has mean 50 and standard deviation of 30, of seeing a deviation as large (in absolute value) as 95?
    c. What is the probability, if the true distribution has mean 0.5 and standard deviation of 0.3, of seeing a deviation as large (in absolute value) as zero?
9. A paper by Chiappori, Levitt, and Groseclose (2002) looked at the strategies of penalty kickers and goalies in soccer. Because of the speed of the play, the kicker and goalie must make their decisions simultaneously (a Nash equilibrium in mixed strategies). For example, if the goalie moves to the left when the kick also goes to the left, the kick scores 63.2% of the time; if the goalie goes left while the kick goes right, then the kick scores 89.5% of the time. In the sample there were 117 occurrences when both players went to the left and 95 when the goalie went left while the kick went right. What is the p-value for a test that the probability of scoring is different? What advice, if any, would you give to kickers, based on these results? Why or why not?
10. A paper by Claudia Goldin and Cecelia Rouse (1997) discusses the fraction of men and women who are hired by major orchestras after auditions. Some orchestras had applicants perform from behind a screen (so that the gender of the applicant was unknown) while other orchestras did not use a screen and so were able to see the gender of the applicant. Their data show that, of 445 women who auditioned from behind a screen, a fraction 0.027 were "hired". Of the 599 women who auditioned without a screen, 0.017 were hired. Assume that these are Bernoulli random variables. Is there a statistically significant difference between the two samples? What is the p-value? Explain the possible significance of this study.
11. Another paper, by Kristin Butcher and Anne Piehl (1998), compared the rates of institutionalization (in jail, prison, or mental hospitals) among immigrants and natives. In 1990, 7.54% of the institutionalized population (or 20,933 in the sample) were immigrants. The standard error of the fraction of institutionalized immigrants is 0.18. What is a 95% confidence interval for the fraction of the entire population who are immigrants? If you know that 10.63% of the general population at the time are immigrants, what conclusions can be made? Explain.
12. Calculate the probability in the following areas under the Standard Normal pdf with mean of zero and standard deviation of one. You might usefully draw pictures as well as making the calculations. For the calculations you can use either a computer or a table.
    a. What is the probability, if the true distribution is a Standard Normal, if seeing a value as large as 1.75?
    b. What is the probability, if the true distribution is a Standard Normal, if seeing a value as large as 2?
    c. If you observe a value of 1.3, what is the probability of observing such an extreme value, if the true distribution were Standard Normal ?
    d. If you observe a value of 2.1, what is the probability of observing such an extreme value, if the true distribution were Standard Normal ?
    e. What are the bounds within which 80% of the probability mass of the Standard Normal lies?
    f. What are the bounds within which 90% of the probability mass of the Standard Normal lies?
    g. What are the bounds within which 95% of the probability mass of the Standard Normal lies?
13. Consider a standard normal pdf with mean of zero and standard deviation of one.
    a. Find the area under the standard normal pdf between -1.75 and 0.
    b. Find the area under the standard normal pdf between 0 and 1.75.
    c. What is the probability of finding a value as large (in absolute value) as 1.75 or larger, if it truly has a standard normal distribution?
    d. What values form a symmetric 90% confidence interval for the standard normal (where symmetric means that the two tails have equal probability)? A 95% confidence interval?
14. Now consider a normal pdf with mean of 3 and standard deviation of 4.
    a. Find the area under the normal pdf between 3 and 7.
    b. Find the area under the normal pdf between 7 and 11.
    c. What is the probability of finding a value as far away from the mean as 7 if it truly has a normal distribution?
15. If a random variable is distributed normally with mean 2 and standard deviation of 3, what is the probability of finding a value as far from the mean as 6.5?
16. If a random variable is distributed normally with mean -2 and standard deviation of 4, what is the probability of finding a value as far from the mean as 0?
17. If a random variable is distributed normally with mean 2 and standard deviation of 3, what values form a symmetric 90% confidence interval?

18. If a random variable is distributed normally with mean 2 and standard deviation of 2, what is a symmetric 95% confidence interval? What is a symmetric 99% confidence interval?
19. A random variable is distributed as a standard normal. (You are encouraged to sketch the PDF in each case.)
    a. What is the probability that we could observe a value as far or farther than 1.7?
    b. What is the probability that we could observe a value nearer than 0.7?
    c. What is the probability that we could observe a value as far or farther than 1.6?
    d. What is the probability that we could observe a value nearer than 1.2?
    e. What value would leave 15% of the probability in the left tail?
    f. What value would leave 10% of the probability in the left tail?
20. A random variable is distributed with mean of 8 and standard deviation of 4. (You are encouraged to sketch the PDF in each case.)
    a. What is the probability that we could observe a value lower than 6?
    b. What is the probability that we could observe a value higher than 12?
    c. What is the probability that we'd observe a value between 6.5 and 7.5?
    d. What is the probability that we'd observe a value between 5.5 and 6.5?
    e. What is the probability that the standardized value lies between 0.5 and -0.5?
21. You know that a random variable has a normal distribution with standard deviation of 16. After 10 draws, the average is -12.
    a. What is the standard error of the average estimate?
    b. If the true mean were -11, what is the probability that we could observe a value between -10.5 and -11.5?
22. You know that a random variable has a normal distribution with standard deviation of 25. After 10 draws, the average is -10.
    a. What is the standard error of the average estimate?
    b. If the true mean were -10, what is the probability that we could observe a value between -10.5 and -9.5?
23. You are consulting for a polling organization. They want to know how many people they need to sample, when predicting the results of the gubernatorial election.
    a. If there were 100 people polled, and the candidates each had 50% of the vote, what is the standard error of the poll?
    b. If there were 200 people polled?
    c. If there were 400 people polled?
    d. If one candidate were ahead with 60% of the vote, what is the standard error of the poll?
    e. They want the poll to be 95% accurate within plus or minus 3 percentage points. How many people do they need to sample?
24. Using the ATUS dataset that we've been using in class, form a comparison of the mean amount of TV time watched by two groups of people (you can define your own groups, based on any of race, ethnicity, gender, age, education, income, or other of your choice).
    a. What are the means for each group? What is the average difference?
    b. What is the standard deviation of each mean? What is the standard error of each mean?
    c. What is a 95% confidence interval for each mean?
    d. Is the difference statistically significant?

25. For a Normal Distribution with mean -2 and standard deviation of 3, what is the area to the right of 0.7?
26. For a Normal Distribution with mean -3 and standard deviation of 3, what is the area to the left of 1.2?
27. For a Normal Distribution with mean -2 and standard deviation of 9, what is the area to the right of -19.1?
28. For a Normal Distribution with mean -6 and standard deviation of 4, what is the area to the left of -8.8?
29. For a Normal Distribution with mean 13 and standard deviation of 4, what is the area to the left of 7?
30. For a Normal Distribution with mean -4 and standard deviation of 1, what is the area to the right of -2.3?
31. For a Normal Distribution with mean -2 and standard deviation of 9, what is the area to the right of -21.8?
32. For a Normal Distribution with mean 14 and standard deviation of 0, what is the area to the left of 14?
33. For a Normal Distribution with mean 1 and standard deviation of 6, what is the area to the right of -2?
34. For a Normal Distribution with mean 5 and standard deviation of 9, what is the area to the left of -4?
35. For a Normal Distribution with mean 5 and standard deviation of 8, what is the area to the left of 21?
36. For a Normal Distribution with mean 13 and standard deviation of 2, what is the area to the right of 9.6?
37. For a Normal Distribution with mean -10 and standard deviation of 5, what is the area in both tails farther from the mean (in absolute value) than -7?
38. For a Normal Distribution with mean -1 and standard deviation of 3, what is the area in both tails farther from the mean (in absolute value) than -7?

39. For a Normal Distribution with mean 3 and standard deviation of 3, what is the area in both tails farther from the mean (in absolute value) than -4.2?
40. For a Normal Distribution with mean 0 and standard deviation of 5, what is the area in both tails farther from the mean (in absolute value) than -9.5?
41. For a Normal Distribution with mean 3 and standard deviation of 9, what is the area in both tails farther from the mean (in absolute value) than -0.6?
42. For a Normal Distribution with mean -7 and standard deviation of 3, what is the area in both tails closer to the mean (in absolute value) than -11.8?
43. For a Normal Distribution with mean -6 and standard deviation of 8, what is the area in both tails closer to the mean (in absolute value) than -5.2?
44. For a Normal Distribution with mean 4 and standard deviation of 3, what is the area in both tails closer to the mean (in absolute value) than -2.3?
45. For a Normal Distribution with mean -9 and standard deviation of 6, what is the area in both tails closer to the mean (in absolute value) than -7.2?
46. For a Normal Distribution with mean -4 and standard deviation of 6, what is the area in both tails closer to the mean (in absolute value) than 9.2?
47. For a Normal Distribution with mean -13 and standard deviation of 5, what is the area in both tails closer to the mean (in absolute value) than -20?
48. For a Normal Distribution with mean 9 and standard deviation of 5, what is the area in both tails closer to the mean (in absolute value) than 20?
49. For a Normal Distribution with mean -7 and standard deviation of 2, what is the area in both tails closer to the mean (in absolute value) than -10?