

Extra hypothesis testing problems

1. The date is June 1, 1995 — the first day of your summer internship with the Metropolitan Transit Authority (sorry, that person next to you got Goldman Sachs). Your boss stops by your spartan little cubicle and gives you your first assignment. She tells you that the upcoming issue of “Long Island Rail Road News Briefs” will have something extra. Besides a listing of on time performance figures (by branch), the pamphlet will also go “high tech” by commenting on statistical aspects of differences in performance. The file `mtaintern.mpj` contains the on time performance figures for the 11 Long Island Rail Road branches in May 1995 and May 1994. Your boss wants to say something about whether on time performance is different between 1994 and 1995, but sadly, she didn’t learn her statistics at the Stern School, and has forgotten what to do. Your job is to help your boss not look foolish. What do these data say about differences in on time performance between May 1994 and May 1995? Carefully state the tests that you are using, the hypotheses that you are testing, the assumptions that you are making, and your conclusions in both a technically correct way and a nontechnical fashion that your boss could understand.
2. A bank wishes to investigate how long it takes customers to do their business (waiting time and time to perform their transactions, which we call serving time). They believe that average serving time is consistent with 30 customers per hour being “processed” through the system (during the midmorning hours). A consulting firm is hired by the bank, which observes midmorning operations for a while. They report back that during the survey, 20 customers were served, with the average serving time being 3 minutes, and sample standard deviation of serving time being 1.5 minutes.
 - (a) The bank wishes to know if these results are consistent with their prior beliefs about the serving process. Carefully state the null and alternative hypotheses that would be appropriate for such a test.
 - (b) What are your conclusions? Justify your answer. What assumptions are you making in performing your test?
3. A *Time Magazine*/CNN poll conducted in May, 1993 of 500 teenagers (16 and 17 year olds) found that 55% of the respondents said that they had had sexual intercourse. Of this group, 61% said that they used birth control every time. Let p represent the true proportion of all 16 and 17 year olds who have had sex and use birth control every time. A prominent televangelist claims that morals in the U.S. are deteriorating rapidly, with the proof being that “More than 30% of all 16 and 17 year olds are having sex and using birth control every time while doing it!” Do the data support this claim? Carefully state the hypotheses you are testing and the test(s) that you are using.
4. *Of Counsel* magazine, a publication aimed at attorneys, periodically conducts surveys of the number of lawyers employed at various New York law firms. The file `lawyers.mpj` contains data for the number of attorneys at each firm based on the surveys in 1988 and 1991. You are interested in knowing whether the typical number of lawyers at firms of this type has changed between time periods. Carefully state the null and alternative hypotheses for this problem. What are your conclusions? Be sure to justify your answers. What assumptions are you making in performing your test(s)? Do they seem valid here?
5. Consider again the file `zagat.mpj` of data from the 1998 Zagat restaurant guide.
 - (a) Is the average Zagat rating for décor significantly different from that for service? Carefully state the hypotheses you are testing and the test(s) that you are using.
 - (b) A tourist reads in her New York City guidebook that meals in New York aren’t as expensive as some people say; in fact, “the average dinner in New York costs only \$35.” Do the Zagat data support this statement? Carefully state the hypotheses you are testing and the test(s) that you are using.
 - (c) It’s reasonable to consider a rating of 20 as being “good.” Consider a randomly chosen Manhattan restaurant; is the probability significantly different from .5 that it will have food rating at least 20? Carefully state the hypotheses you are testing and the test(s) that you are using.
6. The Privacy and American Business and Opinion Research Corp. conducted a survey of 457 Internet users during February 1999. When asked if they have a problem with companies collecting information

on their buying habits and preferences when the information is used to customize offers and services, 393 replied “No.” Does this provide evidence that the true proportion of Internet users who have a problem with companies doing this is less than 20%? Carefully state the hypotheses you are testing and the test(s) that you are using.

7. The 64 teams selected to the NCAA Men’s Basketball tournament are separated into four regions. The teams are seeded within their region from 1 through 16 based on a consensus of team strength. Thus, in the first four rounds of the tournament, each game involves a higher–seeded team playing a lower–seeded team. During the first ten years of the 64–team format, there were 116 games between teams that were seeded within one level of each other (e.g., first versus second seed, fourth versus fifth seed, etc.). The higher–seeded team won 63 of these games. Test the hypothesis that teams that are seeded one level apart are, in truth, exactly evenly matched. If there is an advantage to the higher or lower seed, proceed without assuming that one or the other would have the advantage. Carefully state the hypotheses you are testing and the test(s) that you are using.
8. Consider again the file `correction.mpj` on the last 14 market corrections.
 - (a) Is the average return for small cap stocks during a market correction significantly different from the average return for the S&P Composite during this time period? Carefully state the hypotheses you are testing and the test(s) that you are using.
 - (b) Is the average return for small cap stocks during a market correction significantly different from the average return for mid cap stocks during this time period? Carefully state the hypotheses you are testing and the test(s) that you are using.