Testing – 6.3 t-tests

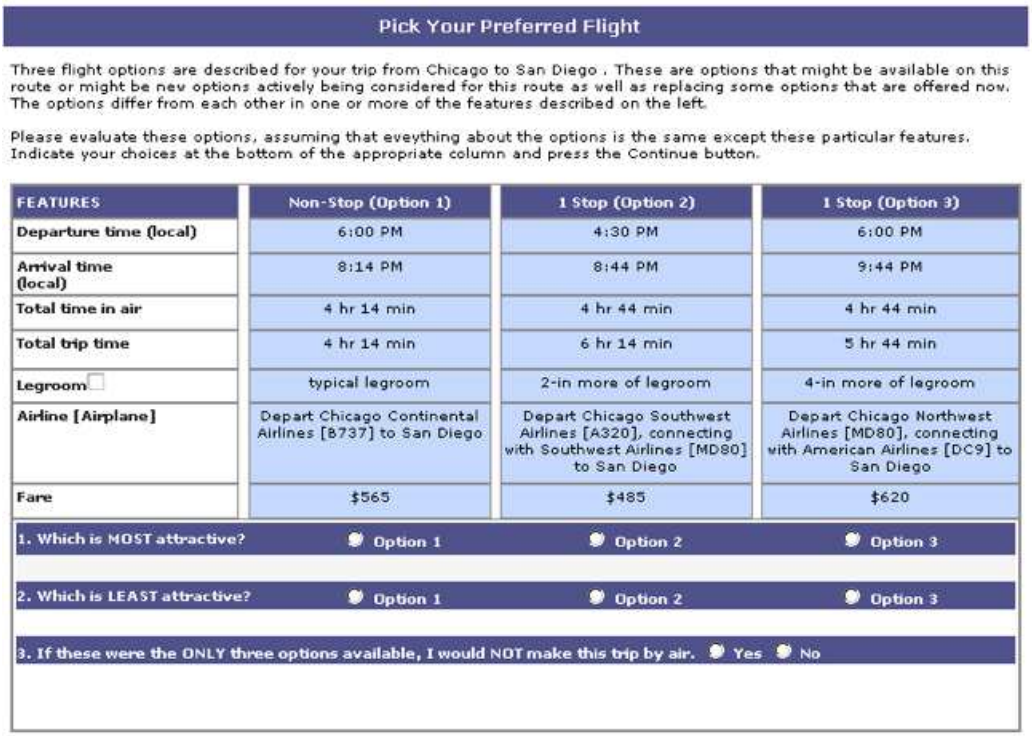
Michel Bierlaire

The Boeing data set

Before we start investigating the methodology for the t-test, we consider another data set that will be used for illustration.

These data come from an Internet choice survey conducted by the Boeing Company in the Fall of 2004. Boeing was interested in understanding the sensitivity that air passengers have toward the attributes of an airline itinerary, such as fare, travel time, transfers, legroom, and aircraft. It was executed on a sample of the customers of an Internet airline booking service. The Internet service takes a specific user request for travel in a city pair and interrogates the web sites of airlines that provide service in that market, returning to the user a compiled list of available itineraries. While that interrogation is taking place, randomly selected customers were recruited to be surveyed.

A typical page of the survey instrument is shown in Figure 1. The respondent was offered three choices based on the origin-destination market request that the respondent entered into the itinerary search engine. The first alternative is always a non-stop flight, the second always a flight with 1 stop on the same airline, and the third is always a flight with 1 stop and a change of airline. The respondent was asked to rank the available choices as well as given the option to decline all of the stated options. Demographic data collected included age, gender, income, occupation, and education. Situational variables that were identified included: a) the desired departure time; b) trip purpose; c) who is paying for the trip; and d) the number in the travel party. All trips were for origin-destination city pairs in the United States.



Pick Your Preferred Flight

Three flight options are described for your trip from Chicago to San Diego . These are options that might be available on this route or might be new options actively being considered for this route as well as replacing some options that are offered now. The options differ from each other in one or more of the features described on the left.

Please evaluate these options, assuming that everything about the options is the same except these particular features. Indicate your choices at the bottom of the appropriate column and press the Continue button.

FEATURES	Non-Stop (Option 1)	1 Stop (Option 2)	1 Stop (Option 3)
Departure time (local)	6:00 PM	4:30 PM	6:00 PM
Arrival time (local)	8:14 PM	8:44 PM	9:44 PM
Total time in air	4 hr 14 min	4 hr 44 min	4 hr 44 min
Total trip time	4 hr 14 min	6 hr 14 min	5 hr 44 min
Legroom	typical legroom	2-in more of legroom	4-in more of legroom
Airline [Airplane]	Depart Chicago Continental Airlines [B737] to San Diego	Depart Chicago Southwest Airlines [A320], connecting with Southwest Airlines [MD80] to San Diego	Depart Chicago Northwest Airlines [MD80], connecting with American Airlines [DC9] to San Diego
Fare	$565	$485	$620

1. Which is MOST attractive? Option 1 Option 2 Option 3

2. Which is LEAST attractive? Option 1 Option 2 Option 3

3. If these were the ONLY three options available, I would NOT make this trip by air. Yes No

Figure 1: The choice of airline itinerary: example of survey instrument

Variable	Description
Subj_ID	Unique identifier for each respondent.
Subj_Male	1 if male, 2 otherwise
Subj_Age	Age, (1 = Less than 18 years, 2 = 18-24 years, 3= 25-34 years, 4 = 35-44 years, 5 = 45-54 years, 6 = 55-64 years, 7 = 65-74 years, 8 = 75 years or older
Subj_Occupation	Occupation (01 = Executive and Managerial, 02 = Professional, 03 = Technicians and related support, 04 = Sales, 05 = Administrative support, 06 = Services, 07 = Precision production, craft, repair, 08 = Machine operators, assemblers, inspectors, 09 = Transportation and material moving, 10 = Handlers, cleaners, helpers, 11 = Farming, forestry, and fishing, 12 = Armed forces)
Subj_Income	Annual income in 100$
Subj_IncomeMissing	Income is missing
Subj_Education	Education (01 = Less than High School Diploma, 02 = High School Graduate, 03 = Some college, No Degree, 04 = Associate Degree - Occupational, 05 = Associate Degree - Academic, 06 = Bachelors Degree, 07 = Masters Degree, 08 = Professional Degree, 09 = Doctorate Degree)

Table 1: The choice of airline itinerary: description of respondent specific variables

Variable	Description
SP1_MostAttractive	SP survey response to "Which is MOST attractive"
SP2_LeastAttractive	SP survey response to "Which is LEAST attractive"
SP3_NotByAir	SP survey response to "If these were the ONLY three options available, I would NOT make this trip by air" (1=yes, 2=no)

Table 2: The choice of airline itinerary: description of survey responses

Variable	Description
Trip_Purpose	Trip purpose (1=business, 2=leisure, 3=attending conference/seminar/training, 4=both business and leisure)
Trip_TravelerPays	1 if the traveler is paying for the trip, 2 if it is his employer, 3 if it is a third party
Trip_IdealDepartureTime	Respondents ideal departure time (hours after midnight)
Trip_PartySize	Number of persons traveling
Trip_OrigMinGMT	Origin city time zone (minutes from GMT (Greenwich Mean Time))
Trip_DestMinGMT	Destination city time zone (minutes from GMT)
Trip_BaseFlightTime	Flight time for shortest non-stop itinerary in minutes
Trip_Miles	Length of itinerary in miles
Trip_Direction	Direction of itinerary (1=East to West, 2=West to East, 3=North-South)

Table 3: The choice of airline itinerary: description of trip specific attributes

Variable	Description
OptX_DepTimeHrs	Option X: Departure time, local (hours after midnight)
OptX_ArrTimeHrs	Option X: Arrival time, local (hours after midnight)
OptX_TotalTimeInAir	Option X: Total time in air (hours)
OptX_TotalTriptime	Option X: Total trip time (hours)
OptX_Legroom	Option X: Legroom in Inches, -2 = 2 inches less than typical, 0 = typical, 2 = 2 inches more than typical, 4 = 4 inches more than typical
OptX_AirlineA	Option X: Airline for first leg (only known to arbitrary airline number for proprietary reasons)
OptX_AirlineB	Option X: Airline for second leg (if there exists a second leg) (only known to arbitrary airline number for proprietary reasons)
OptX_AirplaneA	Option X: Airplane for first leg (only known to arbitrary airplane number for proprietary reasons)
OptX_AirplaneB	Option X: Airplane for second leg (if there exists a second leg) (only known to arbitrary airplane number for proprietary reasons)
OptX_Fare	Option X: Fare ($)
OptX_SchedDelayEarly	Option X: Schedule delay (hours) - early departure (calculated from OptX_DepTimeHrs and Trip_IdealDepartureTime)
OptX_SchedDelayLate	Option X: Schedule delay (hours) - late departure (calculated from OptX_DepTimeHrs and Trip_IdealDepartureTime)

Table 4: The choice of airline itinerary: description of alternative specific attributes where X corresponds to the choice option (1),(2) and (3)

Variable	Average	St. Dev.	Min	Max
Subj_ID	1807.97	1043.03	1.00	3613.00
Subj_Male	0.50	0.50	0	1.00
Subj_Age	3.95	1.14	1.00	8.00
Subj_Occupation	2.54	1.89	1.00	12.00
Subj_Income	107.08	81.40	10.00	350.00
Subj_IncomeMissing	0.1	0.30	0.00	1.00
Subj_Education	5.88	1.71	1.00	9.00
Trip_Purpose	2.04	0.77	1.00	4.00
Trip_TravelerPays	1.20	0.46	1.00	3.00
Trip_IdealDepartureTime	12.75	4.99	0	23.75
Trip_PartySize	1.70	0.99	1.00	5.00
Trip_OrigMinGMT	382.18	82.07	300.00	480.00
Trip_DestMinGMT	397.34	82.86	300.00	480.00
Trip_BaseFlightTime	224.14	95.15	40.00	381.00
Trip_Miles	1568.43	783.79	119.00	2719.00
Trip_Direction	1.91	0.87	1.00	3.00
SP1_MostAttractive	1.45	0.73	1.00	3.00
SP2_LeastAttractive	2.36	0.68	1.00	3.00
SP3_NotByAir	1.60	0.54	1.00	2.00
Opt1_DepTimeHrs	11.72	3.34	6.00	18.00
Opt1_ArrTimeHrs	15.21	3.35	7.67	21.63
Opt1_TotalTimeInAir	3.73	1.59	0.67	6.35
Opt1_TotalTriptime	3.73	1.59	0.67	6.35
Opt1_Legroom	0.92	2.24	-2.00	4.00
Opt1_AirlineA	4.52	2.60	1.00	11.00
Opt1_AirlineB	0.00	0.00	0.00	0.00
Opt1_AirplaneA	4.57	2.30	1.00	8.00
Opt1_AirplaneB	0.00	0.00	0.00	0.00
Opt1_Fare	405.65	199.84	80	1330
Opt1_SchedDelayEarly	2.04	3.98	0.00	17.00
Opt1_SchedDelayLate	2.28	2.91	0.00	21.38

Table 5: The choice of airline itinerary: descriptive statistics of variables

Variable	Average	St. Dev.	Min	Max
Opt2_DepTimeHrs	11.67	3.35	6.00	18.00
Opt2_ArrTimeHrs	16.92	3.36	9.17	24.10
Opt2_TotalTimeInAir	4.23	1.59	1.16	6.85
Opt2_TotalTriptime	5.50	1.67	1.83	8.85
Opt2_Legroom	0.96	2.25	-2.00	4.00
Opt2_AirlineA	4.68	2.67	1.00	11.00
Opt2_AirlineB	0.00	0.00	0.00	0.00
Opt2_AirplaneA	4.46	2.32	1.00	8.00
Opt2_AirplaneB	4.40	2.34	1.00	8.00
Opt2_Fare	407.07	200.93	80.00	1390.00
Opt2_SchedDelayEarly	1.92	4.05	0.00	17.75
Opt2_SchedDelayLate	2.75	2.81	0.00	23.38
Opt3_DepTimeHrs	11.66	3.34	6.00	18.00
Opt3_ArrTimeHrs	16.89	3.41	9.25	24.03
Opt3_TotalTimeInAir	4.24	1.59	1.16	6.85
Opt3_TotalTriptime	5.48	1.67	1.92	8.85
Opt3_Legroom	1.06	2.25	-2.00	4.00
Opt3_AirlineA	4.63	2.61	1.00	11.00
Opt3_AirlineB	4.73	2.67	1.00	11.00
Opt3_AirplaneA	4.49	2.33	1.00	8.00
Opt3_AirplaneB	4.52	2.27	1.00	8.00
Opt3_Fare	405.20	197.65	80.00	1275.00
Opt3_SchedDelayEarly	1.92	3.98	0.00	17.00
Opt3_SchedDelayLate	2.73	2.78	0.00	22.97

Table 6: The choice of airline itinerary: descriptive statistics of variables