Practice Midterm Exam
Statistical Methods and Models - Math 410, Fall 2011
October 21, 2011

You may use a calculator, and you may bring in one sheet (8.5” by 11” or A4) of notes. Otherwise closed book.

(1) (a) You roll a fair (six sided) die twice. The total of the two rolls is 9. Use Bayes’ method to determine the probability that the first roll was a 4.
SOLUTION:*Let E be the event that the first roll is a 4 and F be the event that the total is 9. We are asking for $P(E|F)$*

$$P(E|F) = \frac{P(E\cap F)}{P(F\cap E)+P(F\cap E')} = \frac{1/36}{1/36+3/36} = \frac{1}{4}.$$

(b) You roll a fair (six sided) die twice. Let F be the event that the first roll was a 1. Let E be the event that the total shown is 5. Calculate $P(E|F')$. Is E independent of F?
SOLUTION:

$$P(E|F') = \frac{P(E\cap F')}{P(F')} = \frac{3/36}{5/6} = .1.$$

But $P(E) = 4/36 = 1/9$, so E is not independent of F'. Therefore it is also not independent of F.

(2) Let E and F be possible outcomes of an experiment. Prove that if E is independent of F then F is independent of E. That is, if $P(E) = P(E|F)$ then $P(F) = P(F|E)$.
SOLUTION: *We assume $P(E|F) = P(E)$. We want to check that $P(F) = P(F|E)$.*

$$P(F|E) = \frac{P(F\cap E)}{P(E)}.$$

On the other hand $P(E) = P(E|F) = \frac{P(F\cap E)}{P(F)}$. Substituting this for $P(E)$ in the first equation gives

$$P(F|E) = \frac{P(F\cap E)}{P(F\cap E)/P(F)} = P(F).$$

(3) Suppose you have a screening test for the Lurgy virus. Suppose the prevalence of Lurgy is 1%, the sensitivity of the test is .98 and the specificity is .95. What is the preditictive value positive in this example? What is the predictive value negative?
SOLUTION:

$$PV^+ = P(\textit{infected} \mid \textit{positive}) = \frac{.01\cdot .98}{.01\cdot .98 + .99\cdot(1-.95)} = \frac{.0098}{.0098+.0495} = .1653.$$

$$PV^- = P(\textit{not infected} \mid \textit{negative}\ ) = \frac{.99 * .95}{.99 * .95 + .01 * .02} = \frac{.9405}{.9405+.0002} = .9998.$$

(4) 15 statisticians are giving a training to increase their social competence. They are scored on a 7 point scale before and afterwards with the following results:

Subject :	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Before :	5	3	4	2	1	6	7	3	2	3	5	1	4	4	3
After :	6	2	4	4	3	6	7	5	3	5	5	3	4	5	2

Use a sign test to determine if there sufficient evidence at the .05 signifancance level to conclude that this training increases social competence.

SOLUTION: *There are 15 subjects. 8 display increased social competence after the training, 2 display decreased social competence, and 5 are ties.*

(a) H_0 *is that the training has no effect.*
(b) H_a *is that the training increases social competence.*
(c) *The significance level* $\alpha = .05$.
(d) *The test statistic is that out of the 10 subjects with changed social competence, 8 have increased social competence.*
(e) *The p-value is the probability that* $8, 9$ *or* 10 *of the* 10 *people with changed social competence have an increase, assuming* H_0. *This is:*

$$\binom{10}{10}.5^{10} + \binom{10}{9}.5^{10} + \binom{10}{8}.5^{10} = 56 \cdot .5^{10} = .055.$$

(f) *The p-value is greater than the significance level, so we retain the null hypothesis that the test has no effect.*

(5) You wish to test whether your cat prefers one paw over the other. You dangle a ribbon in front of your cat 10 times, and he bats it with his right paw 8 of those time, and with his left paw two of those times.

Use a binomial exact test to determine if this is sufficient evidence at the .05 significance level to conclude that your cat prefers one paw over the other.

SOLUTION:

(a) H_0 *is that the cat is no more likely to use one paw than the other when batting ribbons.*
(b) H_a *is that the cat is more likely to use one paw than the other when batting ribbons.*
(c) *The significance level* α *is* .05.
(d) *The test statistic is 8 right paws out of 10 trials.*
(e) *The p-value of this, given* H_0 *is the probability of a result at least this extreme. That would be 0, 1, 2, 8, 9 or 10 times using the right paw. The calculation here (out of 10 trials) is just like the previous probles, except we get twice the probability since*

we are looking at a two-tailed example. So we get a p-value of .11.

(f) *We retain the null hypothesis that our cat is not left or right pawed.*

(6) (This is a slight over-simplification of a real study.) In the FUTURE II study of the HPV vaccine, 10,565 womeon were divided randomly into two groups. One group was given Gardasil, and the other group was given a placebo. After three years, 43 women received a diagnosis related to cervical cancer. 1 was in the vaccine group, 42 were in the placebo group.

Is this sufficient evidence to conclude that Gardasil protects against cervical cancer at the .01 significance level? (Use the binomial exact test.)

SOLUTION:

(a) *H_0 is that the vaccine has no effect, and thus the positive diagnoses are equally likely to be in the placebo group and the vaccinated group. ($\pi = .5$ where π is the change of a positive case being in the placebo group.)*

(b) *H_a is that the positive diagnoses are more likely to be in the placebo group. ($\pi > .5$.)*

(c) *The significance level is $\alpha = .01$.*

(d) *The test statistic is that, out of our sample of 43 positive cases, 42 are in the placebo group.*

(e) *The p-value is the probability (assuming H_0) that 42 or more of our sample ended up in the placebo group. This is*

$$\binom{43}{43}(.5)^{43} + \binom{43}{42}(.5)^{43} = 44(.5)^{43} = 5 \times 10^{-12}.$$

(f) *Our p-value is (much) smaller than the significance level, so we reject H_0 and conclude that the vaccinated group is less likely to have a positive diagnosis.*

(7) Let K be a binomial variable with $n = 3$, $\pi = .4$. Calculate the expected value, variance and standard deviation of K.

SOLUTION:

$$E(K) = .6^3 \cdot 0 + 3 \cdot .4 \cdot .6^2 \cdot 1 + 3 \cdot .4^2 \cdot .6 \cdot 2 + .4^3 \cdot 3 = 0 + .432 + .576 + .192 = 1.2.$$

$$E(K^2) = 0 + .432 + 1.152 + .576 = 2.16.$$

So

$$Var(K) = 2.16 - 1.44 = .72$$

So $SD(K) = .8485$.

(8) Let X be a continuous random variable, uniform on $[1, 5]$. Give the pdf and cdf for X and graph them both.

SOLUTION: *The pdf is*

$$f(x) = \begin{cases} 0 & x < 1 \\ .25 & 1 \le x \le 5 \\ 0 & x > 5. \end{cases}$$

The cdf is

$$F(x) = \begin{cases} 0 & x \le 1 \\ .25(x-1) & 1 \le x \le 5 \\ 1 & x \ge 5. \end{cases}$$

(I'm not putting the graphs in but you should.)

(9) One of the 4 functions below is the pdf of a random variable. Which one? (Say briefly what is wrong with the others.)

(a)

$$f(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x \ge 1. \end{cases}$$

(b)

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \le x \le 2 \\ 0 & x \ge 2 \end{cases}$$

(c)

$$f(x) = \begin{cases} 0 & x < -1 \\ x & -1 \le x \le \sqrt{3} \\ 0 & x \ge \sqrt{3} \end{cases}$$

(d)

$$f(x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \le x \le 2 \\ 0 & x \ge 2 \end{cases}$$

SOLUTION: *The first has total integral .5. The second has total integral 2. The third is sometimes negative. The fourth is a pdf, since it is non-negative and has total integral 1.*