Math 216 Exam 1 Solutions

Fall 2008 - Brad Hartlaub

#1. Since the psychologist is interested in detecting any differences in the two populations, we use the two-sided Kolmogorov-Smirnov test statistic.

$$H_0: [\, F(t) = G(t), \text{ for every } t \,]$$

$F$: w/out learning disabilities; $G$: with learning disabilities

$$H_1: [\, F(t) \neq G(t), \text{ for at least one } t \,]$$

| X (without l.d.) $m=7$ | Y (with l.d.) $n=8$ | $\lvert F_m(t) - G_n(t) \rvert$ |
|---|---|---|
| 183 | | $\lvert 1/7 - 0 \rvert = 8/56$ |
| 191 | | $\lvert 2/7 - 0 \rvert = 16/56$ |
| 197 | | $\lvert 3/7 - 0 \rvert = 24/56$ |
| | 202 | $\lvert 3/7 - 1/8 \rvert = 17/56$ |
| 204 | | $\lvert 4/7 - 1/8 \rvert = 25/56$ |
| 218 | | $\lvert 5/7 - 1/8 \rvert = 33/56$ |
| | 220 | $\lvert 5/7 - 2/8 \rvert = 26/56$ |
| 227 | | $\lvert 6/7 - 2/8 \rvert = 34/56$ |
| | 228 | $\lvert 6/7 - 3/8 \rvert = 27/56$ |
| 233 | | $\lvert 1 - 3/8 \rvert = 35/56$ ← Max. |
| | 239 | $\lvert 1 - 4/8 \rvert = 28/56$ |
| | 242 | $\lvert 1 - 5/8 \rvert = 21/56$ |
| | 243 | $\lvert 1 - 6/8 \rvert = 14/56$ |
| | 261 | $\lvert 1 - 7/8 \rvert = 7/56$ |
| | 343 | $\lvert 1 - 1 \rvert = 0/56$ |

$m = 7,\ n = 8 \Rightarrow d = 1$

Thus, $$J = \frac{mn}{d} \max_{i=1,\dots,N} \left\{ \lvert F_m(z_i) - G_n(z_i) \rvert \right\} = \frac{7 \cdot 8}{1} \cdot \frac{35}{56} = 35.$$

Math 216 Exam 1 Solutions
Fall 2008 - Brad Hartlaub.

#1. From Table A.10, the p-value $= P(J \geq 35) = .0559$
Thus, at the $\alpha = .05$ level, we cannot reject $H_0$.
i.e., we do not have statistically significant evidence of any differences between the populations of students with and without learning disabilities.
(Notice that we would make a different conclusion at the $\alpha = .06$ level - or any higher level.)

Dotplots and descriptive statistics clearly show a large outlier (343) for the students with disabilities. Thus, the Miller Jackknife procedure should be used the formally test for differences in the variability for these two populations. If, as suspected, the variances are significantly different, then the Fligner-Policello test based on placements should be used to test for a difference in the two medians.

Math 216 Exam 1 Solutions
Fall 2008 - Brad Hartlaub

#2. Let $\eta_1$ = median response (percentage variation in blood sugar) for dose 1 and $\eta_2$ = median response for dose 2. These two sets of rabbits are independent, so use Stat > nonparametrics > Mann-Whitney. According to Minitab, a 95.8% C.I. for $\eta_1 - \eta_2$ is $(-17.42, -2.40)$. Notice that zero is not in this C.I. Thus, at the $\alpha = .042$ level, we can reject $H_0: \eta_1 = \eta_2$ in favor of $H_1: \eta_1 \neq \eta_2$. That is, we have a statistically significant difference in the median responses for these two doses of a drug.

Explanation: To find the C.I., all 81 differences $X_i$ (Dose 1) - $Y_j$ (Dose 2) were formed. These differences were then sorted from smallest to largest. According to Table A.6, with $\alpha = .02$, $m = 9$, and $n = 9$, $w_{.02} = 109$. Thus,

$$C_\alpha = \frac{n(2m+n+1)}{2} + 1 - w_{\alpha/2}$$

$$= \frac{(9)(18+9+1)}{2} + 1 - 109 = 18.$$

$$\Delta_L = U^{(18)} \text{ and } \Delta_U = U^{(81+2-18)} = U^{(64)}$$

$-17.42$ (verified with Minitab) $-2.40$

Math 216 Exam 1 Solutions

Fall 2008 - Brad Hartlaub

#3. $H_0: p = .10$ vs. $H_1: p = .20$ when $\alpha \le .05$

| | Approximate Power With Minitab (For > alternative) | Approximate Power With Minitab (For ≠ Alternative) |
|---|---|---|
| $n = 50$ | .703373 | .617672 |
| $n = 100$ | .897308 | .848537 |
| $n = 150$ | .966242 | .944298 |
| $n = 200$ | .989329 | .980565 |
| $n = 300$ | .999021 | .997883 |

↳ For our test, the power and sample size program in Minitab indicates the required $n = 102$.

To illustrate the exact procedures, suppose $n = 100$. Using the c.d.f command or Graph > Prob Dist Plot > Binomial we find

| $b$ | 15 | 16 | 17 |
|---|---|---|---|
| $P(B \ge b)$ | .0726 | .0399 | .0206 |

If we select $\alpha = .0399$, then when $p = .2$

Exact Power $= P(B \ge 16) = .871$, which is slightly lower than .897308 above.

Summary Table:

| $n$: | 50 | 100 | 150 | 200 | 300 |
|---|---|---|---|---|---|
| $\alpha$: | .025 | .04 | .044 | .043 | .038 |
| Exact: | .556 | .872 | .963 | .989 | .999 |
| Approx: | | .897 | .966 | .989 | .999 |

Math 216 Exam 1 Solutions
Fall 2008 - Brad Hartlaub

#4. Let $\theta$ = median # of pages in all 114 books.
The distribution of the number of pages is not symmetric (see histogram or dotplot) so use the sign test.

$H_0: \theta = 225$
$H_1: \theta \neq 225$

$B = 9$ Minitab

| N | Below | Equal | Above |
|---|---|---|---|
| 12 | 3 | 0 | 9 |

P-value = .146

Since .146 > .05, we cannot reject $H_0$ at the $\alpha = .05$ level. We do not have evidence to refute the hypothesis that $\theta = 225$.

- A 96.14% C.I. for $\theta$ is (156, 454)
$(Z^{(3)}, Z^{(10)})$

Math 216 Exam 1 Solutions
Fall 2008 - Brad Hartlaub

#5. Let $z_i = Y_i$ (lab 2) $- X_i$ (lab 1) $i = 1, \dots, 10$ be the differences in the counts for the two labs and $\theta$ be the population median difference. This is clearly paired data. This differences are not symmetric, so the sign test & interval should be used.

$H_0: \theta = 0$
$H_1: \theta \neq 0$
$B = 3$

Minitab:

| N | Below | Equal | Above |
|---|---|---|---|
| 10 | 7 | 0 | 3 |

P-value = .3438

Since $.3438 > .05$, we cannot reject $H_0$. There does not appear to be a ~~significant~~ difference in the lab counts for these two labs.

A 97.85% CI for $\theta$ ~~given~~ is: $(-1.2, .3)$, which includes 0.

#6. No, the van der Waerden test should not be used because the data are clearly paired. The van der Waerden test is used for two independent samples.