استنباط آماری

پاییز ۱۴۰۴

تمرین سوم

مرتضی ملکی‌نژاد شوشتری - ۸۱۰۱۰۴۲۵۶

سوال ۱

۱

فرض صفر این است که $\mu_A \geq \mu_B$ و فرض جایگزین این است که $\mu_A < \mu_B$.

این یک تست one tail است و برای بررسی آن باید یک two sampled t test انجام شود.

۲

باید pooled variance حساب شود:

$$s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2} = \frac{9 * 16 + 99 * 625}{108} \approx 574.25 \Rightarrow s_p = \sqrt{574.25} \approx 23.96$$

حال با استفاده از pooled variance می‌توان test statistic را حساب کرد و سپس آن را با یک توزیع t با $n_A + n_B - 2$ درجه آزادی مقایسه کرد:

$$t = \frac{\bar{x}_A - \bar{x}_B}{s_p\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \approx \frac{-6}{7.947} \approx -0.755$$

$$critical\ value : t_{0.05,108} \approx -1.659 \Rightarrow Reject\ if\ h_0 < -1.659$$

$$-0.755 > -1.659 \Rightarrow h_0\ is\ not\ rejected$$

پس یعنی فرض صفر رد نمی‌شود.

۳

داریم:

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s_A^2}{n_1} + \frac{s_B^2}{n_2}}} = \frac{-6}{\sqrt{1.6 + 6.25}} \approx -2.142$$

$$dof = 10 - 9 = 9 \Rightarrow Critical\ value = t_{0.05,9} \approx -1.83$$

چون مقدار test statistic از مقدار critical value کمتر نیست فرض صفر رد نمی‌شود.

۴

در اینجا یکی شد. در حالت کلی تست unpooled چون واریانسی بیشتر از pooled variance نیاز دارد، معمولا critical region کوچکتری دارد و اگر تست pooled رد نشود، احتمال رد شدن این تست کمتر است.

۵

با استفاده از welch t test داریم [۱] :

$$dof = \frac{(\frac{s_A^2}{n_1} + \frac{s_B^2}{n_2})^2}{\frac{(\frac{s_A^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_B^2}{n_2})^2}{n_2 - 1}} = \frac{(1.6 + 6.25)^2}{\frac{2.56}{9} + \frac{39.0625}{99}} \approx 179.10 t_{0.05,179.10} = -1.653$$

یعنی باز هم تست فرض صفر را رد نمی‌کند.

سوال ۲

1. Modeling and parameter of interest

a

$$x = after - before$$

b

پارامتری از جمعیت که این آماره تخمین می‌زند تفاوت مقدار لگاریتم‌ها است. یعنی داریم:

$$x = log(after) - log(before) = log(\frac{after}{before}) \Rightarrow exp(x) = \frac{after}{before}$$

پس یعنی می‌توان با این آماره، نسبت گلوکز قبل و بعد مصرف اندازه گرفت.

در تست‌ها چیزی که اندازه گرفته می‌شود، میانگین این پارامتر است.

2. Hypothesis formulation

a

فرض صفر انتظار دارد که دارو مقدار گلوکز را کاهش ندهد، پس یعنی نسبت میزان گلوکز بعد از مصرف به قبل از مصرف از یک بیشتر باشد پس یعنی مقدار x از صفر بیشتر باشد:

$$H_0 : \bar{x} \geq 0$$
$$H_1 : \bar{x} < 0$$

b

همانطور که بیان شد، اگر دارو تاثیری در کاهش مقدار گلوکز نداشته باشد، داریم:

$$after \geq before \Rightarrow \frac{after}{before} \geq 1 \Rightarrow log(\frac{after}{before}) \geq 0 \Rightarrow log(after) - log(before) \geq 0$$

c

در صورت سوال بیان شد که پس از گرفتن لگاریتم، خاصیت نویزها بصورت جمع در می‌آید، پس یعنی انگار x حاصل جمع تعدادی متغیر تصادفی است. اگر تعداد این متغیرها زیاد باشد (که درموضوعی مثل میزان قندخون، عوامل زیادی دخیل هستند و این فرض بعید نیست.) طبق قضیه حد مرکزی، توزیع x از یک توزیع نرمال تبعیت می‌کند، پس t test مناسب است. از آنجایی که تعداد نمونه‌ها به نسبت کم است باید از توزیع t استفاده کرد. و میانگین مقادیر x را چک کرد.

3. Inference

a

$$\bar{x} = \frac{1}{15}\sum(After - Before) = 0.06 \qquad s \approx 0.046$$
$$statistic = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \approx -5.09$$

b

$$dof = n - 1 = 14 \quad p_{value} = P(t_{14} < t) \approx 0.00008$$

همانطور که حساب شد، مقدار p value با اختلاف از مقدار α کمتر است یعنی تست فرض صفر را رد می‌کند.

4. Interval estimation

a

$$\bar{x} = -0.06 \quad s = 004 \qquad CI = [\bar{x} - t_{14,0.95} * \frac{s}{\sqrt{n}}, \bar{x} + t_{14,0.95} * \frac{s}{\sqrt{n}}] \Rightarrow CI_{0.95} \approx [-0.08, -0.03]$$

b

با به توان رساندن داریم:

$$exp \Rightarrow e^{CI} \approx [0.92, 0.96]$$

یعنی به احتمال ۹۵ درصد، میانگین کاهش گلوکز چیزی حدود چهار تا هشت درصد است.

5. Diagnostic reasoning

a

با این‌کار، داده‌ها از فضای پیوسته به فضای گسسته می‌آیند و نمی‌توان درباره پراکندگی و یا مقدار تاثیر صحبتی کرد. همچنین بین داده‌ای که تنها ۰۱.۰ کمتر شده و داده‌ای که ۰۹.۰ کمتر شده فرقی نیست. استفاده از همچنین چیزی در شرایطی که داده‌های بهتری داریم منطقی نیست. در عمل اینکار شبیه به sign test است و در جاهایی که توزیع و مقادیر معلوم هستند، تست‌های پارامتریک از این تست بهتر هستند.

b

پراکندگی و میزان تاثیر دارو از دست می‌روند و در کیفیت تست به اینصورت اثر دارند که تست حاصل شده، از power کمتری برخوردار خواهد بود. همچنین دیگر نمی‌توان درباره میزان تاثیر دارو صحبتی کرد.

سوال ۳

۱

این یک matched test است. در هر گروه باید میزان بهبود را حساب کرد و سپس میانگین و واریانس نمونه‌ای حساب شود و تست روی آن انجام شود.

$$diff = drug\ change - placebo\ change \qquad t = \frac{\bar{diff}}{\frac{\sigma}{\sqrt{n}}}$$

با انجام محاسبات، مقدار p value در هرگروه بدست می‌آید:

جدول ۱: مقدار p value در هر دسته

Warden	P Value
A	0.11
B	0.22

یعنی در هر دو گروه، تست t با confidence level برابر با ۹۵.۰ نمی‌تواند فرض صفر را رد کند.

سوال ۴

با توجه به اینکه در اینجا مسئله تعداد اعضای هر دسته است می‌توان از chi squared goodness of fit استفاده کرد. داریم:

$$\sum_{i=1}^{k} \frac{O_i - E_i}{E_i} \sim \chi^2_{k-1}$$

پس در اینجا داریم:

$$\sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{k} \frac{(N_i - 80)^2}{80} = \sum_{i=1}^{k} \frac{N_i^2 - 160N_i + 6400}{80} = \sum_{i=1}^{k} \frac{N_i^2}{80} + (80 - 2N_i) = \sum_{i=1}^{k} \frac{N_i^2}{80} + 400 - 2 *$$

$$400 = \sum_{i=1}^{k} \frac{N_i^2}{80} - 400 \sim \chi^2_4$$

از آنجایی که تست با confidence level برابر با ۰۱.۰ است، برای critical value باید مقدار ppf مربوط به $\chi_{0.99,4}$ را حساب کرد که تقریبا برابر است با 13.277 پس داریم:

$$\sum_{i=1}^{k} \frac{N_i^2}{80} - 400 > 13.277 \Rightarrow \sum_{i=1}^{k} \frac{N_i^2}{80} > 413.277 \Rightarrow \sum_{i=1}^{k} N_i^2 > 33061.6$$

سوال ۶

۱

تست از نوع one sample و two tail است. مقدار significance level برابر است با $P(reject\ H_0|H_0)$ یعنی :

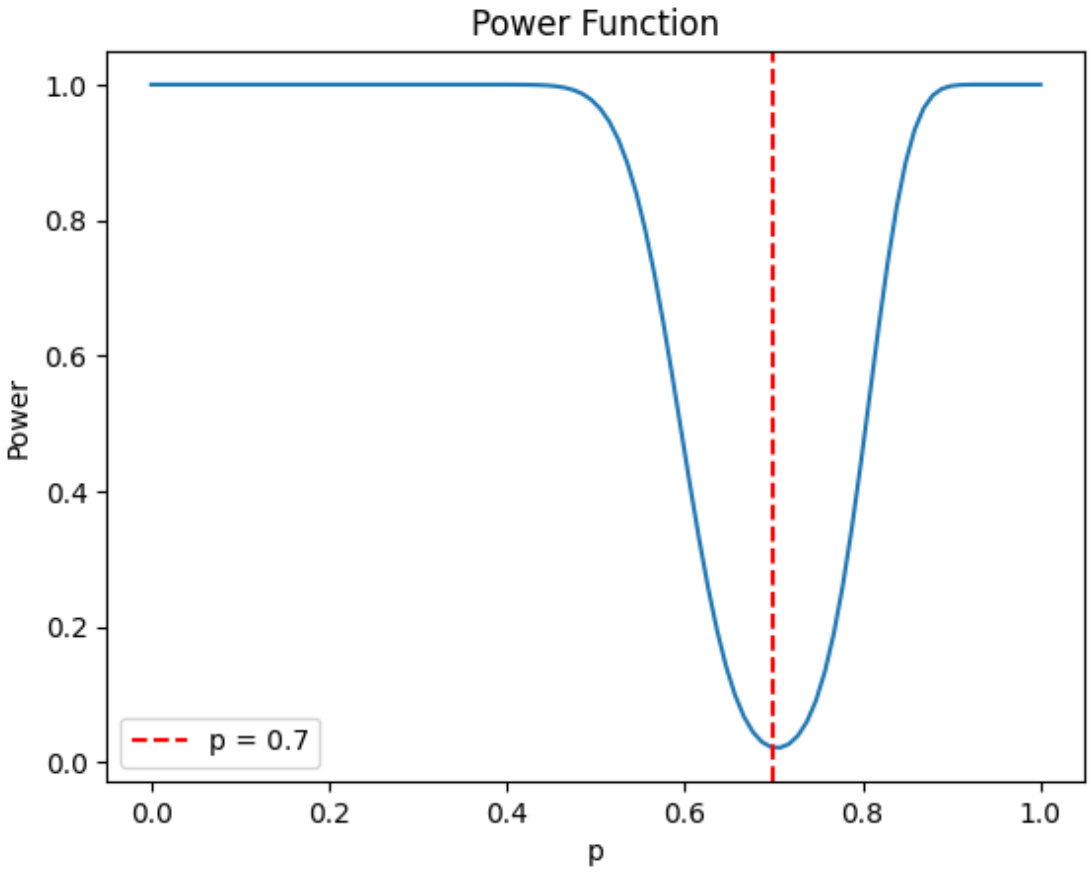
$$P(X < 60|p = 0.7) + P(X > 80|p = 0.7)$$

با استفاده از فرمول PMF توزیع دوجمله‌ای داریم:

$$P(X = x) = C(n, x)(p^x)(1 - p)^{n-x} = C(100, x) * 0.7^x * 0.3^{100-}$$

با محاسبه مقدار $P(X < 60|p = 0.7) + P(X > 80|p = 0.7)$ به این می‌رسیم: $P(X < 60|p = 0.7) + P(X > 80|p = 0.7) \approx 0.021$ یعنی مقدار α تقریبا برابر 0.021 است.

۲



شکل ۱: مقدار power به ازای مقادیر مختلف p

۳

a

$$Critical\ value = P(|X-70| \geq 2|p=0.7) = 1 - P(X=69) - P(X=70) - P(X=71) \approx 0.74$$

اگر مقدار α از این مقدار بیشتر باشد فرض صفر رد می‌شود ولی در هیچ تست منطقی‌ای مقدار α اینقدر زیاد نیست.

b

داریم:

$$\mu = np = 70 \quad \sigma = \sqrt{np(1-p)} \approx 4.58 \Rightarrow z = \frac{S + correction - \mu}{\sigma} \Rightarrow Z_{correction} = \frac{3}{4.58} \approx 0.55 \quad z_{without\ correction} = \frac{2.5}{4.58} \approx 0.66$$

حال برای محاسبه critical value داریم:

$$P(|Z| > z_{crrection}) \approx 2 * 0.29 = 0.58$$

$$P(|Z| > z_{without\ crrection}) \approx 2 * 0.29 = 0.5$$

سوال ۷

۱

دراینجا باید از آزمون chi squared goodness of fit استفاده کرد.

برای محاسبه مقدار expect می‌توان از pdf توزیع دوجمله‌ای استفاده کرد. طبق داده‌های سوال داریم:

جدول ۲: مقدار p value در هر دسته

Number of boys	Expect	Ovserve
0	16	26
1	48	32
2	48	40
3	16	30

سوال عملا مقدار p value را می‌خواهد پس داریم:

$$\sum_i = 0^3 \frac{(O_i - E_i)^2}{E_i} \approx 25.16$$
$$p_value = 1 - cdf_3(25.16) \approx 1.42e - 5$$

پس یعنی به‌ازای $\alpha < 1.42e - 5$ تست فرض صفر را رد نمی‌کند. و به‌ازای سایر مقادیر فرض صفر رد می‌شود.

۲

دراینجا باید از ump استفاده کرد و شبیه قسمت قبل عمل کرد، صررفا بجای استفاده از یک p مشخص از p مربوط به maximum likelihood استفاده کرد:

$$P(x) = C(3,x) * (p^x) * (1-p)^{3-x} \Rightarrow P_{total} = \prod_{i=0}^{3} C(3,i) * (p^i) * (1-p)^{3-i} = 1*3*3*1*p^6*(1-p)^6 = 9*(p-p^2)^6$$

اگر مقدار قدرمطلق $p - p^2$ بیشینه باشد، مقدار likelihood بیشینه می‌شود (به شرط اینکه $0 < p < 1$). پس یعنی در $p = 0.5$ تابع بیشینه می‌شود. پس یعنی عملا نتایج قسمت قبل که با $p = 0.5$ حساب شدند،همان نتایج مربوط به UMP هستند. و در همان $\alpha < 1.42e-5$ فرض صفر رد نمی‌شود. و به‌ازای سایر مقادیر فرض صفر رد می‌شود.

سوال ۱۰

۱

a

این یک تست one sampled برای میانه است و می‌توان از آزمونی مثل sign test برای آن استفاده کرد.

H_0 : median ≥ 120

H_1 : median < 120

b

داریم:

$data-120 = [5, -10, 10, -2, 20, 0, -15, 15, 2, -1, 25, 30, -5, 8, 1, 18, -8, 0] \Rightarrow t = \text{number of negatives} = 4 \quad n = 16$

برای مقادیری که روی ۱۲۰ هستند، آن‌ها را از تست خارج می‌کنیم و تست را با $n^* = 16$ ادامه می‌دهیم.

۲

a

برای محاسبه p value با استفاده از توزیع دوجمله‌ای داریم:

$$P(x) = C(16, x)(0.5)^{16} \Rightarrow P(x \leq 6) = \frac{1}{2^{16}} \sum_{i=0}^{6} C(16, i) = \frac{14893}{2^{16}} \approx 0.2272$$

b

از آنجایی که p value از significance level بزرگ‌تر است، تست فرض صفر را رد نمی‌کند.

۳

a

داریم:

$$\mu = n * p = 16 * 0.5 = 8 \quad \sigma^2 = n * p * (1-p) = \frac{n}{4} = 4 \Rightarrow \sigma = 2z = \frac{t + \text{continuity correction} - \mu}{\sigma} = \frac{-1.5}{2} = -0.75$$

b

با مراجعه به جدول توزیع نرمال استاندارد داریم $CDF_Z(-0.75) \approx 0.2266$ که تقریبا با جواب مرحله قبل برابر است و تست همچنان نمی‌تواند فرض صفر را رد کند.

References

[1] Wikipedia contributors. *Welch's t-test*. https://en.wikipedia.org/wiki/Welch's_t-test. Accessed: 2024-02-04.