روش‌های ناپارامتری

کتاب دکتر جواد بهبودیان

دانشگاه اراک

۱۳۹۹

صفحات: ۱۸۶–۱۹۰

ارائه دهنده:

محراب عتیقی

تمارین فصل یازدهم

۱- سه پسر و چهار دختر، به تصادف در یک صف می‌ایستند. جدول تابع احتمال دوها را مستقیماً و از طریق توابع احتمال‌گردش‌ها پیداکنید.

پاسخ:

تعداد جایگشت‌ها برابراست با $\frac{7!}{3!4!} = 35$

جایگشت	R_1	R_2	R
xxxyyyy	1	1	2
xxyxyyy	2	2	4
xxyyxyy	2	2	4
...	...	...	...

تمارین فصل یازدهم

ادامه تمرین 1:

R_1	1	2	3
احتمال	$\frac{1}{7}$	$\frac{4}{7}$	$\frac{2}{7}$

R_2	1	2	3	4
احتمال	$\frac{4}{35}$	$\frac{18}{35}$	$\frac{12}{35}$	$\frac{1}{35}$

R	2	3	4	5	6	7
احتمال	$\frac{2}{35}$	$\frac{5}{35}$	$\frac{12}{35}$	$\frac{9}{35}$	$\frac{6}{35}$	$\frac{1}{35}$

تمارین فصل یازدهم

۲- تابع احتمال R_1, R_2, R، را باشمارش و بدون استفاده از f_{R_1,R_2} پیداکنید.

پاسخ: چگالی R_1 را ازراه شمارش با روش زیر می‌یابیم:

با n_1 مهره از نوع x وبا n_2 مهره از نوع y می‌توانیم $\binom{n_1+n_2}{n_1}$ جایگشت بسازیم. حال تعداد جایگشتهایی را می‌یابیم که برای آنها $R_1 = r_1$ برای این منظور از n_1 مهره از نوع x_1 را در ظرف r_1 میریزیم به‌طوری که هیچ یک جای خالی نماند. این کار را به $\binom{n_1-1}{r_1-1}$ طریق می‌توانیم انجام دهیم.

تمارین فصل یازدهم

ادامه تمرین ۲)

حال n_2مهره از نوع y را روی یک خط میچینیم و r_1 ظرف را در فاصله آنها قرار می‌دهیم. چون n_2+1 فاصله داریم، این کار را به $\binom{n_2+1}{r_1}$ طریق می‌توان انجام داد. پس

$$P(R_1=r_1)=\frac{\binom{n_1-1}{r_1-1}\binom{n_2+1}{r_1}}{\binom{n_1+n_2}{n_1}}$$

به همین ترتیب برای چگالی R_2 داریم:

$$P(R_2=r_2)=\frac{\binom{n_2-1}{r_2-1}\binom{n_1+1}{r_2}}{\binom{n_1+n_2}{n_1}}$$

تمارین فصل یازدهم

۳–در تمرین ۱ مقدار **P(R=5)** را مستقیماً، از روی تابع احتمال **R**، و با تقریب نرمال محاسبه کنید.

مستقیماً

$$P(R=5)=\frac{9}{35}\approx 0.257$$

بااستفاده ازتقریب نرمال:

$$\mu_R=1+\frac{2n_1n_2}{n_1+n_2}=1+\frac{24}{7}=\frac{31}{7}\approx 4.43$$

$$\sigma_R^2=\frac{(\mu_R-1)(\mu_R-2)}{n_1+n_2-1}=1.39$$

$$P(R=5)=P(4.5<R<5.5)$$

$$=P\left(\frac{4.5-4.43}{1.18}<\frac{R-\mu_R}{\sigma_R}<\frac{5.5-4.43}{1.18}\right)$$

$$\approx \phi(0.91)-\phi(0.06)\approx 0.2947$$

تمارین فصل یازدهم

٤- ثابت کنید که ماکزیمم R برابراست با

$$2min(n_1,n_2) + 1 - \delta_{n_1,n_2}$$

بطوری که

$$\delta_{n_1,n_2} = \begin{cases} 1 & n_1 = n_2 \\ 0 & n_1 \neq n_2 \end{cases}$$

پاسخ:ماکزیمم R زمانی رخ می‌دهد که n_1چیز از نوع x, n_2چیز از نوع y یک درمیان قرار گرفته باشند که سه‌حالت زیررا درنظر می‌گیریم:

الف- با گردش‌های نوع xشروع و پایان یابد(ماکزیمم زمانی است که $n_1 = n_2 + 1$ پس $R = 2n_2 + 1$ از طرفی $min(n_1,n_2) = n_2$

بنابراین $max(R) = 2min(n_1,n_2) + 1$

تمارین فصل یازدهم

ب- با گردش‌های نوع y شروع و پایان یابد(ماکزیمم زمانی است که $n_2 = n_1 + 1$ پس $R = 2n_1 + 1$ از آنجا که $min(n_1,n_2) = n_1$ پس $maxR = 2min(n_1,n_2) + 1$).

ج- بادوهای مختلف شروع و پایان یابد(ماکزیمم زمانی است که:
$n_1 = n_2$ و $R = 2n_1$ چون $min(n_1,n_2) = n_1$
بنابراین $max(R) = 2min(n_1,n_2)$

بادرنظر گرفتن حالات فوق به راحتی دیده می‌شود که ماکزیمم R برابر است با

$$2min(n_1, n_2) + 1 - \delta_{n_1,n_2}$$

$$\delta_{n_1,n_2} = \begin{cases} 1 & n_1 = n_2 \\ 0 & n_1 \neq n_2 \end{cases}$$

تمارین فصل یازدهم

۵– کواریانس R_1,R_2 را پیداکنید.

پاسخ:

$$Var(R)=Var(R_1+R_2)=Var(R_1)+Var(R_2)+2Cov(R_1,R_2)$$

$$Cov(R_1,R_2)=\frac{1}{2}\left[Var(R)-\left(Var(R_1)-Var(R_2)\right)\right]$$

$$=\frac{1}{2}\left[\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}-\frac{n_{1_2}[(n_1-1)(n_2+1)+(n_2-1)(n_1+1)]}{(n_1+n_2)^2(n_1+n_2-1)}\right]$$

$$=\frac{n_1n_2(n_2-1)(n_1-1)}{(n_1+n_2-1)(n_{1+}n_2)}$$

تمارين فصل يازدهم

۶- درتمرين ۱ ميانگين و واريانس و ضريب همبستگی دوهای R_1,R_2 رامحاسبه كنيد.

پاسخ:می‌دانیم که $E[R_1]=\frac{n_1(n_2+1)}{n_1+n_2}$

$$Var(R_1)=\frac{n_1n_2(n_1-1)(n_2+1)}{(n_1+n_2)^2\,(n_1+n_2-1)}$$

و با تعويض n_1,n_2می‌توانیم $Var(R_2),E[R_2]$ راهم پيداكنيم، ولذا داريم:

$$E[R_1]=\frac{15}{7}\ ,\qquad E[R_2]=\frac{16}{7}$$

$$Var(R_1)=\frac{20}{49}\ ,\qquad Var(R_2)=\frac{24}{49}$$

$$Cov(R_1,R_2)=\frac{216}{49}$$

$$\rho(R_1,R_2)=\frac{Cov(R_1,R_2)}{\sqrt{Var(R_1)*Var(R_1)}}=\frac{54}{\sqrt{30}}\approx 9.859$$

تمارین فصل یازدهم

۷- یک تست درست- نادرست به ترتیب زیر پاسخ داده‌شده است:

$$TFFTFTFTTFTFFTFTFTTF$$

آیا این تست به تصادف پاسخ داده‌شده است؟ (با میزان $\alpha = 0.1$).

پاسخ:می‌دانیم که $\mu_R = 1 + \frac{2n_1n_2}{n_1+n_2}$ و $\sigma_R^2 = \frac{(\mu_R-1)(\mu_R-2)}{n_1+n_2-1}$

پس داریم:

$$\mu_R = 11\,, \sigma_R = 2.18\,, r = 16(n_1 + n_2)$$

$$p - value = P(R \geq 16) \approx P(R \geq 15.5)$$

$$\approx 1 - \phi\left(\frac{15.5 - 11}{2.18}\right) = 0.0197 < 0.1$$

پس، درنتیجه فرض تصادفی بودن را به‌ازای $\alpha = 0.1$، رد می‌کنیم.

تمارین فصل یازدهم

۸- دندان پزشکی می‌خواهد خمیردندان **A,B** را مقایسه کند. به پنج نفر **A** و به سه نفر **B** را توصیه می‌کند. بعد از یک‌سال کرم‌خوردگیها را می‌شمارد و نتایج را بدست می‌آورد:

A:1,6,0,3,5   B:4,2,7

آیا این دو خمیردندان هم اثر می‌باشند؟ (با میزان $\alpha = 0.05$).

پاسخ: ابتدا دونمونه را درکنار یکدیگر به ترتیب صعودی مرتب می‌کنیم:

$$0, 1, 2, 3, 4, 5, 6, 7$$

واضح هست که، $n_1 = 3, n_2 = 3$، ولذا داریم:

$$\mu_R = 4.75 \ , \sigma_R = 1.214 \ , r = 6$$

$$p - value = P(R \geq 6) \approx P(R \geq 5.5) \approx 1 - \phi(0.62) = 0.2676 > 0.05$$

با توجه به مقدار پی، لزومی بر ردکردن فرضیه مربوط به اینکه دونمونه ازیک توزیع، آمده‌اند، نداریم.

تمارین فصل یازدهم

۹- ده نفر پسر و دختر هفت ساله را که درشرایطی تقریباًمساوی بزرگ شده‌اند وزن می‌کنیم و نتایج زیر را بدست می‌آویم(برحسب کیلوگرم).

وزن پسرها: ۳۶و۳۲و۴۰و۲۸و۳۱

وزن دخترها: ۲۹و۳۰و۳۳و۲۷

آیاوزن‌پسرهاودخترهای هفت‌ساله هم‌توزیع می‌باشند؟($\alpha = 0.1$).

پاسخ: $27\,28\,29\,30\,31\,32\,33\,36\,40$

واضح هست که $n_1 = 3$,$n_2 = 3$ ولذا داریم:

(براساس مرتب شده صعودی) $r = 6$, $\sigma_R = 1.38$, $\mu_R = 5.44$

$$p - value = P(R \geq 6) \approx P(R \geq 5.5)$$
$$\approx 1 - \phi(0.04) \approx 0.484 > 0.1$$

به‌ازای $\alpha = 0.1$ فرض هم‌توزیع بودن وزن پسرهاودخترهارا ردمی‌کنیم.

تمارین فصل یازدهم

۱۰- تعداد دانشجویانی را که ساعت ۸ صبح ۲۰ روزتحصیلی، یک اتوبوس از خوابگاه به دانشگاه می‌برد، اعداد زیر می‌باشند:
۲۱و۲٤و۲۳و۱۹و۳۲و۲۸و۲٦و۱۷و۲۰و۲۸و۳۰و۲٤و۱۳و۳۵و۲٦و۲۱و۱۹و ۲۹و۲۷و۱۸

آیا این داده‌ها یک نمونه تصادفی هستند؟ ($\alpha = 0.01$).

$$median = 24$$

پاسخ: 01001110011101100110

$$r = 11 \quad , n_1 = 5 \quad , n_2 = 6$$

$$\mu_R = 10.9 \quad , \sigma_R = 2.153$$

باتوجه‌به‌نکات مربوطبه بخش ۱۱-۲-۲-(تصادفی بودن داده‌های‌عددی)

داریم‌که: ناحیه بحرانی برای این آزمون می‌شود $R \notin (10,17)$به ازای $\alpha = 0.01$، فرض متساوی بودن داده‌ها رد نمی‌شود.

تمارین فصل یازدهم

۱۱- عده‌ای مرد و زن دریک صف، برای خرید شیر، به ترتیب زیر منتظر می‌باشند

$$WMWWMMMWWWWMMMMMWMWM$$

آیا به تصادف ایستاده‌اند؟($\alpha = 0.1$).

پاسخ:بعد از مرتب کردن و صعودی کردن داده‌ها داریم:

$$r = 10 \quad , n_1 = 5 \quad , n_2 = 5$$

$$\mu_R = 10.9 \quad , \sigma_R = 2.153$$

باتوجه‌به‌نکات مربوطبه بخش ۱۱-۲-۲-(تصادفی بودن داده‌های‌عددی)

داریم‌که: ناحیه بحرانی برای این آزمون می‌شود $R \notin (10,17)$به ازای $\alpha = 0.01$، فرض تصادفی بودن ایستادن زنان‌ومردان ردمی‌شود.