فصل ۱۲ آمار ناپارامتری

۱۲-۱ مقدمه

روش‌هایی که در فصل‌های قبل برای استنباط آماری مطرح شدند بر فرض‌هایی در مورد جامعه استوار بودند. در بیشتر موارد فرض بر این بود که جوامعی که نمونه‌ها از آنها استخراج شده‌اند، دارای توزیع نرمال هستند. گاهی چنین فرض‌هایی صادق نبوده و لذا استنباط حاصل، اعتبار کافی را نخواهد داشت. در این موارد می‌توان از روش‌های ناپارامتری که در این فصل مورد بحث قرار می‌گیرند، استفاده نمود.

در روش‌های ناپارامتری هیچ فرضی در مورد توزیع جوامع یا پارامترهای آنها از جمله میانگین و واریانس وجود ندارد. روش‌های ناپارامتری نسبت به روش‌های پارامتری از اطلاعات کمتری استفاده می‌کنند و از این رو در مبحث آزمون فرض‌ها به آزمون‌هایی با توان کمتر و در مبحث فواصل اطمینان به فواصل بزرگتر با درجه اطمینان کمتر منجر می‌شوند.

برای مثال اگر فاصله اطمینان میانگین حاصل از قضیه حد مرکزی با اطمینان ۹۵٪ به صورت (۱۰ و ۲۰) باشد در حالت ناپارامتری با همان اطمینان ۹۵٪، این فاصله طولانی‌تر خواهد بود (مثلاً ۵ و ۲۵). در روش‌های ناپارامتری از همه اطلاعات موجود در داده‌ها استفاده نشده و اغلب به جای خود داده‌ها از رتبه آنها استفاده می‌شود. مثلاً به جای ارتفاع چهار گیاه که بر حسب سانتی‌متر اندازه‌گیری و بصورت ۱۳/۵، ۱۵/۰، ۲۲/۲، و ۴۵/۳ ثبت شده، ممکن است از رتبه آنها یعنی ۱، ۲، ۳ و ۴ استفاده شود. در نتیجه تفاوت ارتفاع دو گیاه اول برابر با تفاوت ارتفاع دو گیاه آخر فرض شده و به تفاوت واقعی بین داده‌ها توجه نمی‌شود. آزمون‌های ناپارامتری غالباً بر روی صفات کیفی چون رنگ گل یا میزان خسارت و موارد مشابه به کار می‌رود، که در آن مشاهدات را می‌توان به صورت رتبه‌هایی یادداشت نمود. موارد زیر عمده‌ترین آزمون‌های ناپارامتری هستند. برای هر مورد، آزمون پارامتری متناظر با آن درون پرانتز آمده است:

۱- آزمون علامت (آزمون t بر روی یک نمونه یا نمونه‌های جفت شده)

۲- آزمون رتبه علامت‌دار ویلکاکسون (آزمون t بر روی یک نمونه)

۳- آزمون نمونه‌های جفت شده ویلکاکسون (آزمون t بر روی نمونه‌های جفت شده)

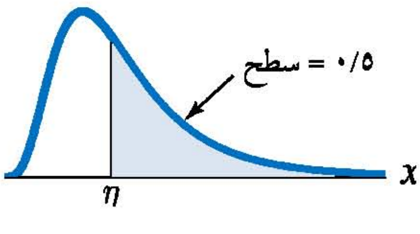
٤- آزمون من‌ویتنی (آزمون t بر روی دو نمونه مستقل)

۵- آزمون کروسکال والیس (آزمون F در تحلیل واریانس یک‌طرفه)

٦- آزمون فریدمن (آزمون F در طرح بلوک‌های کامل تصادفی یا تحلیل واریانس دوطرفه)

۱۲-۲ آزمون علامت

بر خلاف آزمون‌های پارامتری z یا t که در مورد میانگین جامعه (μ) به کار می‌روند، آزمون علامت در مورد میانه جامعه (η) به کار می‌رود. میانه صدک پنجاهم جامعه بوده و کمتر تحت تأثیر چولگی توزیع و داده‌های پرت قرار می‌گیرد.



نمودار ۱۲-۱. موقعیت میانه در یک جامعه دارای چولگی.

در آزمون علامت فرض صفر، برابری میانه جامعه با مقداری ثابت بصورت $H_0 : \eta = \eta_0$ است و فرض مقابل بسته به مورد می‌تواند یک‌طرفه یا دوطرفه باشد. هیچ شرطی در مورد شکل توزیع جامعه وجود ندارد و در صورتی که داده‌ها دارای چولگی شدید باشند، می‌توان آن را به عنوان جایگزینی برای آزمون z (برای نمونه‌های بزرگ) یا t (برای نمونه‌های کوچک) مد نظر قرار داد. برای انجام آزمون در یک نمونه تصادفی، به داده‌های بزرگتر از میانه علامت + و به داده‌های کوچکتر از آن علامت − تعلق گرفته و داده‌های برابر با میانه به حساب نمی‌آیند. توجه کنید که تحت فرض صفر تعداد علائم مثبت یا منفی از یک توزیع دوجمله‌ای پیروی می‌کند و در

صورتی که تعداد علائم مثبت و منفی تفاوت معنی‌داری با هم داشته باشند، فرض صفر رد خواهد شد.

مثال ۱۲-۱. در تحقیقی ۸ دوقلوی یکسان انتخاب شده‌اند. همه این کودکان با هم در یک سن وارد کلاس اول شده‌اند با این تفاوت که از هر دوقلو، یکی از آنها در سال گذشته دوره کودکستان را طی کرده است. جدول زیر نمرات آنها را در پایان کلاس اول نشان می‌دهد. آیا در سطح ۵٪ می‌توان گفت که طی کردن دوره کودکستان نمرات کودکان را در کلاس اول افزایش داده است؟

کودکستان رفته	۶۴	۸۱	۶۷	۷۰	۶۴	۶۷	۷۴	۸۳
کودکستان نرفته	۶۵	۷۷	۶۳	۶۸	۶۶	۶۳	۷۴	۷۸

برای حل مسئله ابتدا باید تفاوت‌های مثبت و منفی را در ردیفی جداگانه یادداشت کرد.

کودکستان رفته	۶۴	۸۱	۶۷	۷۰	۶۴	۶۷	۷۴	۸۳
کودکستان نرفته	۶۵	۷۷	۶۳	۶۸	۶۶	۶۳	۷۴	۷۸
تفاوت‌ها	۱-	۴+	۴+	۲+	۲-	۴+	۰	۵+

اگر گذراندن دوره کودکستان در افزایش نمره بی‌تأثیر باشد، تعداد علائم مثبت و منفی نباید تفاوت معنی‌داری با هم داشته باشند، به عبارتی میانه تفاوت‌ها باید صفر باشد. لذا فرض‌های تحقیق را می‌توان به صورت

$$H_0 : \eta = 0$$
$$H_1 : \eta > 0$$

نوشت. در اینجا ۵ علامت مثبت و ۲ علامت منفی وجود دارد. از آنجا که آزمون، یک‌طرفه راست است، تعداد علائم مثبت بزرگتر از صفر به عنوان آماره آزمون در نظر گرفته شده و با S نشان داده می‌شود. تحت فرض صفر، تعداد علائم مثبت و منفی مشاهده شده از توزیع دوجمله‌ای تبعیت می‌کند. اگر H_0 درست باشد، احتمال مشاهده تفاوت‌های بزرگتر و کوچکتر از صفر با هم برابر بوده و p در توزیع دوجمله‌ای مربوطه برابر با ۰/۵ خواهد بود. n نیز برابر با مجموع تعداد علائم مثبت و منفی است.

احتمال مشاهده حداقل ۵ علامت مثبت در توزیع دوجمله‌ای مذکور با $n = 7$ برابر با

$$P(x \geq ۵) = ۰/۲۲۷$$

است که نحوه محاسبه آن در فصل سوم (بند ۳-۶) توضیح داده شده است. این احتمال در واقع همان مقدار احتمال در آزمون یک‌طرفه مورد بررسی است و در صورت کوچکتر بودن آن از مقدار $\alpha = ۰/۰۵$ فرض صفر رد می‌شود. پس در اینجا تعداد علائم مثبت و منفی تفاوت معنی‌داری با هم نداشته و طی کردن دوره کودکستان، نمرات کودکان را افزایش نداده است.

آزمون علامت: برابری میانه جامعه با یک مقدار خاص

فرض صفر	فرض مقابل
$H_0 : \eta = \eta_0$	$H_1 : \eta > \eta_0$ (آزمون یک‌طرفه راست)
	$H_1 : \eta < \eta_0$ (آزمون یک‌طرفه چپ)
	$H_1 : \eta \neq \eta_0$ (آزمون دوطرفه)
آماره آزمون (S):	تعداد داده‌های بزرگتر از η_0 (آزمون یک‌طرفه راست)
	تعداد داده‌های کوچکتر از η_0 (آزمون یک‌طرفه چپ)
	ماکسیمم تعداد داده‌های بزرگتر یا کوچکتر از η_0 (آزمون دوطرفه)
مقدار احتمال:	$P(x \geq S)$ (در آزمون یک‌طرفه)
	$۲P(x \geq S)$ (در آزمون دوطرفه)

x دارای توزیع دوجمله‌ای با پارامترهای n و $p = ۰/۵$ است و در صورتی که مقدار احتمال کمتر از α باشد فرض صفر رد می‌شود. در آزمون میانه، نمونه باید به طور تصادفی از یک جامعه دارای توزیع احتمال پیوسته انتخاب شوند.

جدول ۱۲-۱. خروجی برنامه Minitab برای مثال ۱۲-۱.

```
Sign test of median =  0.00000 versus > 0.00000

    N  Below  Equal  Above        P  Median
C1  8      2      1      5   0.2266   3.000
```

مثال ۱۲-۲. برای شانزده مشاهده ۱۲، ۱٤، ۲۵، ۳۱، ۱۱، ۱٦، ۵۲، ۱۹، ۹، ۲۵، ۱۵، ۲۱، ۱۸، ۱۲، ۱۵، ٤۰، فرض صفر مبنی بر $\eta = ۲۵$ را در مقابل $\eta \neq ۲۵$ در سطح $\alpha = ۰/۰۵$ آزمون کنید.

حل: پس از قرار دادن علائم مثبت و منفی برای داده‌ها داریم:

۱۲	۱٤	۲۵	۳۱	۱۱	۱٦	۵۲	۱۹	۹	۲۵	۱۵	۲۱	۱۸	۱۲	۱۵	٤۰
−	−		+	−	−	+	−	−		−	−	−	−	−	+

مشاهده می‌شود که سه عدد بزرگتر از ۲۵ و یازده عدد کوچکتر از ۲۵ هستند. لـذا صـرف‌نظـر از تعداد اعداد مساوی با ۲۵، n برابر با ۱٤ خواهد بود. در این آزمون دو طرفه، مقدار احتمال برابـر است با دو برابر احتمال مشاهده ۱۱ منفی یا بیشتر در توزیع دوجمله‌ای با n=۱٤ و p = ۰/۵ کـه برابر

$$p - value = ۲P(x \geq ۱۱) = ۲ \times ۰/۰۲۸۷ = ۰/۰۵۷٤$$

است. از آنجا که مقدار احتمال از مقدار α بیشتر است، لذا میانه نمونه مورد بررسی (۱۷) تفاوت معنی‌داری با ۲۵ ندارد. البته با توجه به اینکه توزیع دوجمله‌ای با p = ۰/۵ متقارن است، می‌تـوان مقدار احتمال مربوط به توزیع دوجمله‌ای را با در نظر گرفتن تصحیح پیوستگی، از طریق تقریب توزیع نرمال نیز به دست آورد. احتمال مشاهده ۱۱ منفی یا بیشتر در توزیع دوجمله‌ای بـا n =۱٤ و p = ۰/۵ از طریق تقریب توزیع نرمال برابر

$$P(x \geq ۱۱) = P(z \geq \frac{(S - ۰/۵) - np}{\sqrt{npq}}) = (z \geq \frac{(۱۰/۵) - ۱٤ \times ۰/۵}{\sqrt{۱٤ \times ۰/۵ \times ۰/۵}}) = P(z \geq ۱/۸۷) = ۰/۰۳۰۷$$

اسـت. بـا توجـه بـه اینکـه آزمـون دو طرفـه اسـت، مقـدار احتمـال برابـر اسـت بـا $۲P(z \geq ۱/۸۷) = ۰/۰٦۱٤$. این مقدار به مقدار احتمال واقعی (۰/۰۵۷٤) خیلی نزدیـک اسـت و بـا توجه به اینکه از مقدار α بیشتر است، فرض صفر رد نخواهد شد. البته روش سـاده‌تـر، مقایسـه آماره آزمون محاسبه شده (۱/۸۷) با $z_{\alpha/۲} = z_{۰/۰۲۵}$ یعنی ۱/۹٦ است. در اینصورت نیـز بـه دلیـل کوچکتر بودن آماره آزمون، دلیل کافی برای رد فرض صفر وجود ندارد.

جدول ۱۲-۲. حل مثال ۱۲-۲ در برنامه Minitab و خروجی مربوطه.

مراحل:

۱. داده‌های را در ستون اول صفحه داده‌ها وارد کنید.

۲. وارد مسیر زیر شوید:

Stat > Nonparametrics > 1-Sample-Sign…

۳. ستون اول را انتخاب کرده و در مقابل گزینه Test median عدد ۲۵ را وارد کنید.

۴. در مقابل گزینه Alternative نیز فرض مقابل (در این مثال: not equal) را انتخاب کرده و بر روی OK کلیک کنید.

```
Sign test of median = 25.00 versus not = 25.00

     N  Below  Equal  Above        P  Median
C1  16     11      2      3   0.0574   17.00
```

آزمون علامت در مورد میانه با استفاده از تقریب توزیع نرمال

فرض صفر: $H_0 : \eta = \eta_0$

فرض مقابل:

$H_1 : \eta > \eta_0$ (آزمون یک‌طرفه راست)

$H_1 : \eta < \eta_0$ (آزمون یک‌طرفه چپ)

$H_1 : \eta \neq \eta_0$ (آزمون دوطرفه)

آماره آزمون

$$z_0 = \frac{(S - 0/5) - np}{\sqrt{npq}}$$

S برابر است با تعداد داده‌های بزرگتر از η_0 در آزمون یک‌طرفه راست یا تعداد داده‌های کوچکتر از η_0 در آزمون یک‌طرفه چپ یا ماکسیمم تعداد داده‌های بزرگتر و کوچکتر از η_0 در آزمون دوطرفه، n برابر است با مجموع تعداد داده‌های بزرگتر یا کوچکتر از η_0 و p نیز برابر است با ۰/۵.

نواحی رد:

$z_0 \geq z_\alpha$ (آزمون یک‌طرفه راست)

$z_0 \leq -z_\alpha$ (آزمون یک‌طرفه چپ)

$|\, z_0 \,| \geq z_{\alpha/2}$ (آزمون دوطرفه)

۱۲-۳ آزمون رتبه علامت‌دار ویلکاکسون

این آزمون همانند آزمون علامت، میانه جامعه را مورد آزمون قرار می‌دهد ولی چون علاوه بر علامت، رتبه‌ها را نیز مد نظر قرار می‌دهد از آن پرتوان‌تر است. در انجام این آزمون فرض بر این است که توزیع جامعه متقارن بوده ولی الزاماً نرمال نیست. در صورتی که توزیع داده‌ها دارای چولگی شدید باشد، بهتر است از طریق تبدیل داده‌ها آنها را تقریباً متقارن نمود. مراحل انجام این آزمون به صورت زیر است:

۱. داده‌های برابر با میانه را نادیده بگیرید.

۲. تفاضل داده‌ها را از میانه به دست آورده و قدر مطلق تفاضل‌های حاصل را از کوچک به بزرگ رتبه‌بندی کنید. برای تفاضل‌های برابر، از رتبه‌های گره‌خورده استفاده کنید (مثال ۱۲-۳).

۳. مجموع رتبه‌های مربوط به تفاضل‌های مثبت $(T+)$ و مجموع رتبه‌های مربوط به تفاضل‌های منفی $(T-)$ را جداگانه به دست آورید و کوچکترین آنها را با T نشان دهید.

٤. دقت کنید که تساوی $(T+)+(T-)=n(n+۱)\,/\,۲$ برقرار باشد. در صورت صحت فرض صفر انتظار می‌رود که $T+$ یا $T-$ برابر با $n(n+۱)\,/\,٤$ باشد.

اگر n مساوی یا کمتر از ٤۰ باشد از جدول ضمیمه ٦ برای آزمون فرض استفاده کنید. در آزمون دو طرفه، مقادیر T مساوی یا کوچکتر از عدد جدول منجر به رد فرض صفر می‌شود. در آزمون یک طرفه راست، $T-$ مساوی یا کوچکتر از عدد جدول منجر به رد فرض صفر و در آزمون یک طرفه چپ، $T+$ مساوی یا کوچکتر از عدد جدول منجر به رد فرض صفر خواهد شد. در مورد نمونه‌های بزرگتر از ٤۰ تقریب توزیع نرمال مورد استفاده قرار می‌گیرد. زیرا در این حالت تحت فرض صفر، T دارای توزیع تقریباً نرمال با میانگین $n(n+۱)\,/\,٤$ و واریانس $n(n+۱)(۲n+۱)\,/\,۲٤$ خواهد بود. لذا آماره آزمون برابر

$$z=\frac{|\,T-n(n+۱)\,/\,٤\,|-۰/۵}{\sqrt{n(n+۱)(۲n+۱)\,/\,۲٤}}$$

است. عبارت ۰/۵ در صورت کسر برای تصحیح پیوستگی اضافه شده است زیرا T دارای توزیع گسسته می‌باشد. z محاسبه شده در آزمون دو طرفه در سطح $\alpha=۰/۰۵$ با ۱/۹٦ و در آزمون یک طرفه با ۱/٦٤ مقایسه شده و در صورتی که بزرگتر از آن بود فرض صفر رد می‌شود.

مثال ۱۲-۳. فرض کنید که میانگین عملکرد یک رقم استاندارد ۲ تن در هکتار است. عملکرد ۱۰ کرت کشت شده با یک رقم جدید در همان ناحیه ۲/۶، ۲/۲، ۲/۴، ۲/۴، ۱/۹، ۲/۳، ۱/۷، ۲/۰، ۲/۵ و ۲/۷ تن در هکتار بوده است. آیا می‌توان گفت که عملکرد واریته جدید از واریته استاندارد موجود بیشتر است؟

اگر فرض نرمال بودن داده‌ها قابل پذیرش نباشد، می‌توان از آزمون رتبه علامت‌دار برای مقایسه استفاده کرد. در این حالت فرض صفر این است که میانه جامعه (نه میانگین) برابر با ۲ است.

عملکرد	۲/۶	۲/۲	۲/۴	۲/۴	۱/۹	۲/۳	۱/۷	۲	۲/۵	۲/۷
$x-\eta$	۰/۶	۰/۲	۰/۴	۰/۴	−۰/۱	۰/۳	−۰/۳		۰/۵	۰/۷
$\lvert x-\eta \rvert$	۰/۶	۰/۲	۰/۴	۰/۴	۰/۱	۰/۳	۰/۳		۰/۵	۰/۷
رتبه	۸	۲	۵/۵	۵/۵	۱	۳/۵	۳/۵		۷	۹

چون یکی از داده‌ها برابر با ۲ است، درمحاسبات در نظر گرفته نشده و لذا تعداد داده‌ها به ۹ تقلیل می‌یابد. دقت کنید که برای دو عدد مشابه میانگین دو رتبه مربوطه قرار داده می‌شود. با توجه به اینکه قدر مطلق تفاضل هر کدام از دو داده ۲/۳ و ۱/۷ از میانه برابر با ۰/۳ است، میانگین رتبه‌های ۳ و ٤ یعنی ۳/۵ به هر دوی آنها تعلق می‌گیرد. این گونه رتبه‌ها را رتبه‌های گره خورده می‌نامند. به همین ترتیب رتبه گره‌خورده ۵/۵ به اعداد ۲/٤ و ۲/٤ اختصاص یافته است.

مجموع رتبه‌های مربوط به تفاضل‌های منفی برابر با ٤/۵=۳/۵+۱ =$(T-)$ و مجموع رتبه‌های مربوط به تفاضل‌های مثبت برابر با ٤۰/۵=۹+۷+۳/۵+۵/۵+۵/۵+۲+۸ =$(T+)$ است. مقدار بحرانی T در جدول ضمیمه ٦ برای آزمون دو طرفه مربوط به سطح ۵٪ و ۹ = n برابر با ۵ می‌باشد. با توجه به اینکه مینیمم مقادیر $(T-)$ و $(T+)$ یعنی ٤/۵ کمتر از ۵ است، فرض صفر رد شده و میانه جامعه برابر با ۲ نیست.

جدول ۱۲-۳. حل مثال ۱۲-۳ و خروجی مربوطه در برنامه Minitab.

مراحل:

۱. داده‌های را در ستون اول صفحه داده‌ها وارد کنید.

۲. وارد مسیر زیر شوید:

Stat > Nonparametrics > 1-Sample-Wilcoxon…

۳. ستون اول را انتخاب کرده و در مقابل گزینه Test median عدد ۲ را وارد کنید.

۴. در مقابل گزینه Alternative نیز فرض مقابل (در این مثال: not equal) را انتخاب کرده و بر روی Ok کلیک کنید.

```
Test of median = 2.000 versus median not = 2.000

        N for   Wilcoxon         Estimated
     N   Test  Statistic      P     Median
C1  10      9       40.5  0.038      2.300
```

مقدار احتمال در این آزمون برابر با ۰/۰۳۸ به دست آمده است، در حالی که مقدار احتمال در آزمون رتبه بر روی همین داده‌ها برابر با ۰/۱۸ خواهد شد. لذا می‌توان گفت که آزمون رتبه علامت‌دار نسبت به آزمون علامت از توان بیشتری برخوردار است. اعمال آزمون t یک طرفه نیز بر روی داده‌های اخیر، مقدار احتمالی برابر با ۰/۰۲۶ به دست خواهد داد. پس آزمون t از هر دو آزمون علامت و رتبه علامت‌دار قوی‌تر است ولی اعتبار آن مشروط بر نرمال بودن توزیع داده‌های جامعه است. در حالت کلی همیشه باید از آزمون پرتوان‌تر استفاده نمود البته در صورتی که فرض‌های لازم برای آن آزمون برقرار باشد.

مثال زیر در مورد آزمون رتبه علامت‌دار بر روی نمونه‌های جفت شده می‌باشد که در آن تفاضل‌ها رتبه‌گذاری می‌شوند.

مثال ۱۲-٤. تعداد ۱۰ بوته گوجه‌فرنگی سالم با دو نژاد قارچی تلقیح شدند. نژاد A بر روی برگ دوم هر بوته و نژاد B روی برگ سوم همان بوته اعمال شد. بعد از سپری شدن زمان لازم، تعداد کلنی‌های قارچ بر روی برگ‌های تلقیح شده شمرده و در جدول زیر ثبت شده است. آیا می‌توان گفت که میانه تعداد کلنی‌های قارچ حاصل از دو نژاد با هم متفاوت است؟

بوته	تعداد کلنی A	تعداد کلنی B	A-B	|A-B|	رتبه	رتبه علامت‌دار
۱	۱۱	۹	۲	۲	۳/۵	۳/۵
۲	۱۲	۱٦	−۳	۳	۵	−۵
۳	۵	۹	−٤	٤	٦	−٦
٤	۷	٦	۱	۱	۱/۵	۱/۵
۵	۵	۱٤	−۹	۹	۷	−۷
٦	۵	۱٦	−۱۱	۱۱	۹	−۹
۷	۸	۸	۰	۰	−	−
۸	۹	۱۰	−۱	۱	۱/۵	−۱/۵
۹	۱٦	۱٤	۲	۲	۳/۵	۳/۵
۱۰	۸	۱۸	−۱۰	۱۰	۸	−۸

ابتدا باید قدر مطلق تفاضل‌ها به دست آید (ستون پنجم از راست) و سپس این تفاضل‌ها صرفنظر از مقادیر صفر رتبه‌گذاری شوند. در نهایت به این رتبه‌ها، علامت تفاوت مربوطه به

ترتیبی که در ستون چهارم آمده است تعلق گرفته و در ستون آخر نوشته می‌شود. حال می‌توان آزمون رتبه علامت‌دار را انجام داد. در اینجا مجموع رتبه‌های منفی برابر با $(T-)$ =۳٦/۵ و مجموع رتبه‌های مثبت برابر با ۸/۵ = $(T+)$ است. فرض صفر دلالت بر برابر بودن میانه تعداد کلنی‌ها برای دو نژاد و فرض مقابل دلالت بر متفاوت بودن آنها دارد. مقدار بحرانی T جدول در آزمون دو طرفه برای سطح ۵٪ و ۹ = n برابر با ۵ می‌باشد. با توجه به اینکه مینیمم مقادیر $(T-)$ و $(T+)$ یعنی ۸/۵ بیشتر از ۵ است، فرض صفر رد نمی‌شود. گرچه تعداد اعضای نمونه کم است ولی روش دیگر آزمون فرض در این مثال استفاده از تقریب نرمال به صورت

$$z=\frac{|\,T-n(n+۱)\,/\,٤\,|-۰/۵}{\sqrt{n(n+۱)(۲n+۱)\,/\,۲٤}}=\frac{|\,۸/۵-۲۲/۵\,|-۰/۵}{\sqrt{۹(۱۰)(۱۹)\,/\,۲٤}}=۱/٦$$

است. با توجه به اینکه آماره z به دست آمده کوچکتر از ۱/۹٦ است، دلیل کافی برای رد فرض صفر وجود ندارد. برنامه Minitab تقریب توزیع نرمال را برای آزمون دوطرفه رتبه علامت‌دار به کار می‌برد که نتیجه آن در زیر آمده است.

جدول ۱۲-٤. حل مثال ۱۲-٤ و خروجی مربوطه در برنامه Minitab.

مراحل:

۱. داده‌های نمونه اول و دوم را بترتیب در ستون‌های اول و دوم وارد کنید.

۲. تفاوت داده‌های نمونه اول و دوم را از طریق Calc>Caculator به دست آورده و در ستون سوم ذخیره کنید.

۳. وارد مسیر زیر شوید:

Stat > Nonparametrics > 1-Sample-Wilcoxon...

۴. ستون سوم را انتخاب کرده و در مقابل گزینه Test median عدد ۰ را وارد کنید.

۵. در مقابل گزینه Alternative نیز فرض مقابل (در این مثال: not equal) را انتخاب کرده و بر روی OK کلیک کنید.

```
Test of median = 0.000000 versus median not = 0.000000

        N for   Wilcoxon          Estimated
     N   Test  Statistic       P     Median
C3  10      9        8.5   0.110     -3.500
```

۱۲-٤ آزمون من‌ویتنی برای مقایسه دو جامعه مستقل

مقایسه میانگین دو جامعه با استفاده از آزمون t مستلزم برقراری شرطهایی از جمله نرمال بودن توزیع جوامع مربوطه و برابری واریانس آنها می‌باشد. اگر توزیع جوامع غیر نرمال یا واریانس‌ها متفاوت باشد، می‌توان از آزمون من-ویتنی برای مقایسه میانه جوامع استفاده نمود.

مثال ۱۲-۵. گفته می‌شود که ترتیب سؤالات امتحانی در توان دادن پاسخ درست به سؤالات مؤثر است. برای بررسی این موضوع معلمی دو نمونه تصادفی ۷ نفری و ٦ نفری از دانش‌آموزان را با دو مجموعه سؤال می‌آزماید. سؤالات فقط در ترتیب متفاوت بوده به گونه‌ای که مجموعه سؤالات A از ساده‌ترین به مشکل‌ترین و مجموعه سؤالات B به صورت برعکس مرتب شده‌اند. جدول زیر نمرات دانش‌آموزان را در هر مورد نشان می‌دهد.

سؤالات A	۹۰	۷۸	۸۳	۸۲	۷٥	۹۱	٦٥
سؤالات B	٦٦	۷۸	٥۰	٦۸	٦۲	٦۰	

آیا در سطح $\alpha = ۰/۰٥$ می‌توان گفت که موقعیت توزیع احتمال نمرات یا به عبارتی موقعیت میانه‌ها در دو حالت متفاوت است؟

تجربه نشان داده است که معمولاً توزیع احتمال نمرات دانش‌آموزان نرمال نیست. لذا استفاده از آزمون t ممکن است معتبر نباشد. در مقایسه میانه‌ها با استفاده از آزمون من-ویتنی ابتدا داده‌های هر دو نمونه را به عنوان یک مجموعه واحد در نظر گرفته و آنها را رتبه‌بندی می‌کنیم به طوری که به کوچکترین عدد، رتبه ۱ و به بزرگترین عدد، رتبه ۱۳ تعلق می‌گیرد. از آنجا که دو عدد ۷۸ وجود دارد، رتبه‌های آنها گره خورده است.

سؤالات A	٦٦	۷۸	٥۰	٦۸	٦۲	٦۰	
رتبه	٥	۸/٥	۱	٦	۳	۲	
سؤالات B	۹۰	۷۸	۸۳	۸۲	۷٥	۹۱	٦٥
رتبه	۱۲	۸/٥	۱۱	۱۰	۷	۱۳	٤

اگر توان پاسخ‌دهی به سؤالات متفاوت باشد رتبه‌های بزرگتر در یک گروه و رتبه‌های کوچکتر در گروه دیگر قرار خواهند گرفت. به عبارتی مجموع رتبه‌های یک گروه از مجموع رتبه‌های گروه دیگر بیشتر خواهد بود. در اینجا مجموع رتبه‌های سؤالات A برابر با $R_1 = ۲٥/٥$، مجموع

رتبه‌های سؤالات B برابر با $R_2 = ۶۵/۵$ و مجموع کل رتبه‌ها برابر با $n(n+۱)/۲$ است کـه در آن $n = n_۱ + n_۲$. این رابطه برای کنترل صحت رتبه‌گذاری کاربرد دارد. در مرحله بعد آماره‌های $U_۱$ و $U_۲$ به صورت

$$U_۱ = R_۱ - \frac{n_۱(n_۱+۱)}{۲} = ۲۵/۵ - \frac{۶(۷)}{۲} = ۴/۵,$$
$$U_۲ = R_۲ - \frac{n_۲(n_۲+۱)}{۲} = ۶۵/۵ - \frac{۷(۸)}{۲} = ۳۷/۵$$

محاسبه می‌شود که بین آنها رابطه $U_۱ + U_۲ = n_۱n_۲$ برقرار است. برای آزمون فرض صفر مبنـی بر برابر بودن میانه‌های دو جامعه، مینیمم $U_۱$ و $U_۲$ با عدد بحرانی در جـدول ضـمیمه ۷ مقایسـه می‌شود. در صورتی که مینیمم $U_۱$ و $U_۲$ مساوی یا کمتر از عدد بحرانی جدول مربوط بـه سـطح مورد نظر و حجم نمونه $n_۱$ و $n_۲$ باشد، فرض صفر رد خواهد شد. اگر انجام آزمون دو طرفـه در سطح $\alpha = ۰/۰۵$ مد نظر باشد، کمینه $U_۱$ و $U_۲$ با عدد بحرانی جدول مربوط به سطح $\alpha = ۰/۰۲۵$ مقایسه می‌شود.

n_1	n_2	٪۵	٪۲/۵	٪۱	٪۰/۵	n_1	n_2	٪۵	٪۲/۵	٪۱	٪۰/۵
۲	۴	–	–	–	–	۴	۱۶	۱۴	۱۱	۷	۵
۲	۵	۰	–	–	–	۴	۱۷	۱۵	۱۱	۸	۶
۲	۶	۰	–	–	–	۴	۱۸	۱۶	۱۲	۹	۶
۲	۷	۰	–	–	–	۴	۱۹	۱۷	۱۳	۹	۷
۲	۸	۱	۰	–	–	۴	۲۰	۱۸	۱۴	۱۰	۸
۲	۹	۱	۰	–	–	۵	۵	۴	۲	۱	۰
۲	۱۰	۱	۰	–	–	۵	۶	۵	۳	۲	۱
۲	۱۱	۱	۰	–	–	۵	۷	۶	۵	۳	۱
۲	۱۲	۲	۱	–	–	۵	۸	۸	۶	۴	۲
۲	۱۳	۲	۱	۰	–	۵	۹	۹	۷	۵	۳
۲	۱۴	۳	۱	۰	–	۵	۱۰	۱۱	۸	۶	۴
۲	۱۵	۳	۱	۰	–	۵	۱۱	۱۲	۹	۷	۵
۲	۱۶	۳	۱	۰	–	۵	۱۲	۱۳	۱۱	۸	۶
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.

نمودار ۱۲–۲. قسمتی از جدول مقادیر بحرانی U در آزمون من–ویتنی.

در آزمون دو طرفه اخیر، مینیمم $U_۱$ و $U_۲$ برابر با ۴/۵ است که از مقـدار U مربـوط بـه سـطح ٪۲/۵ یعنی ۶ کمتر است و لذا فرض رد می‌شود.

آزمون من‌ویتنی برای مقایسه دو جامعه مستقل

فرض صفر	فرض مقابل
$H_0 : \eta_1 = \eta_2$	$H_1 : \eta_1 > \eta_2$ (آزمون یک‌طرفه راست)
	$H_1 : \eta_1 < \eta_2$ (آزمون یک‌طرفه چپ)
	$H_1 : \eta_1 \neq \eta_2$ (آزمون دوطرفه)

آماره آزمون: مینیمم U_1 و U_2

مقادیر U_1 و U_2 از رابط $U_1 = R_1 - \frac{n_1(n_1+1)}{2}$ و $U_2 = R_2 - \frac{n_2(n_2+1)}{2}$ به دست می‌آیند که در آنها R_1 برابر با جمع رتبه‌های نمونه اول و R_2 برابر با جمع رتبه‌های نمونه دوم است.

نواحی رد:

مینیمم U_1 و U_2 مساوی یا کمتر از عدد جدول U مربوط به سطح $\alpha / 2$ (آزمون دوطرفه)

U_2 مساوی یا کمتر از عدد جدول مربوط به سطح α (آزمون یک‌طرفه راست)

U_1 مساوی یا کمتر از عدد جدول مربوط به سطح α (آزمون یک‌طرفه چپ)

فرض‌ها:

نمونه‌ها بطور تصادفی و مستقل از هم استخراج شده‌اند و توزیع احتمال جوامعی که نمونه‌ها از آن استخراج شده‌اند پیوسته می‌باشد.

در صورتی که n_1 و n_2 هر دو بزرگ باشند تحت فرض صفر مبنی بر برابری میانه جوامع، هر دوی U_1 و U_2 دارای توزیع تقریباً نرمال با میانگین $n_1 n_2 / 2$ و واریانس $n_1 n_2 (n_1 + n_2 + 1) / 12$ بوده و آزمون را می‌توان به کمک آماره

$$z = \frac{|\, U - n_1 n_2 / 2 \,| - 0/5}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12}}$$

انجام داد که در آن U می‌تواند هر کدام از مقادیر U_1 یا U_2 باشد. آماره z به دست آمده با عدد جدول z مقایسه می‌شود. تصحیح ۰/۵- در صورت به این خاطر است که آماره U توزیع گسسته دارد.

جدول ۱۲-۵ خروجی برنامه Minitab برای این مثال است که در آن نتیجه آزمون در قالب مقدار احتمال نمایش داده شده است. اگر تعداد رتبه‌های گره‌خورده زیاد باشد تصحیح دیگری باید اعمال گردد که در اینجا به آن اشاره نشده است. برنامه Minitab آزمون من-ویتنی را با استفاده از تقریب توزیع نرمال ولی با در نظر گرفتن تصحیح مربوط به رتبه‌های گره‌خورده انجام می‌دهد (جدول ۱۲-۵).

جدول ۱۲-۵. حل مثال ۱۲-۵ و خروجی مربوطه در برنامه Minitab.

مراحل:

۱. داده‌های نمونه اول و دوم را در ستون‌های اول و دوم صفحه داده‌ها وارد کنید.

۲. وارد مسیر زیر شوید.

Stat > Nonparametrics > Mann-Whitney...

۳. در مقابل گزینه First Sample ستون اول و در مقابل Second Sample ستون دوم را وارد کرده و سطح اطمینان مورد نظر و نوع آزمون (یک طرفه یا دو طرفه) را انتخاب کنید.

۴. گزینه OK را انتخاب کنید.

```
        N  Median
Test A  6   64.00
Test B  7   82.00

Point estimate for ETA1-ETA2 is -16.00
96.2 Percent CI for ETA1-ETA2 is (-29.00;-4.00)
W = 25.5
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0223
The test is significant at 0.0221 (adjusted for ties)
```

۱۲-۵ آزمون کروسکال-والیس

این آزمون برای مقایسه میانه‌های چند جامعه به کار می‌رود و تعمیم یافته آزمون من-ویتنی است. آزمون کروسکال والیس یک جایگزین ناپارامتری برای تحلیل واریانس یک‌طرفه است و لذا فرض صفر (H_0) در آن برابر بودن میانه‌های جوامع مورد مقایسه در مقابل فرض عدم برابری آنها (H_1) می‌باشد. در این آزمون فرض بر این است که نمونه‌ها به طور تصادفی و مستقل از جوامع انتخاب شده و شکل توزیع جوامعی که نمونه‌ها از آن استخراج می‌شوند پیوسته هستند.

اگر مقایسه k جامعه مد نظر باشد و $i = 1,\cdots k$ و n_i حجم نمونه استخراج شده از جامعه iام را نشان دهد، حجم کل نمونه‌ها برابر با $N = n_1 + \cdots + n_k$ است. برای انجام آزمون ابتدا کل مشاهدات از کوچک به بزرگ رتبه‌بندی می‌شوند. مانند قبل به مشاهدات یکسان، رتبه‌های گره‌خورده (میانگین رتبه‌ها) تعلق می‌گیرد. میانگین رتبه‌ها در هر نمونه به صورت

$$\overline{R}_i = \frac{R_i}{n_i}$$

محاسبه می‌شود، که در آن R نشان‌دهنده مجموع رتبه‌های نمونه iام است. سپس آماره آزمون به صورت

$$H = \frac{12 \sum n_i (\overline{R}_i - \overline{R})^2}{N(N+1)}$$

به دست می‌آید که در آن $\overline{R}$ برابر

$$\frac{\sum R_i}{N}$$

است. البته بعضی آمارادانان در صورتی که رتبه‌های گره خورده وجود داشته باشد، تصحیحی برای آماره مذکور در نظر می‌گیرند. برنامه Minitab نیز H تصحح شده را در آزمون کروسکال-والیس محاسبه می‌کند. در صورت عدم وجود رتبه‌های گره‌خورده، s^2 یا واریانس رتبه‌ها برابر است با $s^2 = N(N+1) \,/\, 12$. آماره H با استفاده از رابطه زیر نیز به سادگی بدست می‌آید.

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$$

تحت فرض صفر، آماره H دارای توزیع کای‌دو با $k-1$ درجه آزادی است، بویژه اگر حجم هر نمونه از ۵ کمتر نباشد. آزمون کای‌دو آزمونی یک طرفه بوده و فرض صفر در آن هنگامی رد می‌شود که H مساوی یا بزرگتر از $\chi^2_{(k-1,\ \alpha)}$ باشد.

مثال ۱۲-۶. تصور بر این است که افزایش درصد پنبه در پارچه مقاومت آن را در برابر کشش افزایش می‌دهد. لذا محققی ۵ نوع نخ ساخته شده حاوی درصدهای مختلف پنبه (۲۰٪، ۲۵٪، ۳۰٪ و ۳۵٪) را در ۵ تکرار از لحاظ مقاومت نسبت به کشش مورد بررسی قرار داده و نتایج زیر را بدست آورده است.

درصد وزن پنبه (تیمارها)			
۲۰	۲۵	۳۰	۳۵
۱۲	۱۴	۱۹	۷
۱۷	۱۸	۲۵	۱۰
۱۲	۱۸	۲۲	۱۱
۱۸	۱۹	۱۹	۱۶
۱۸	۱۹	۲۳	۱۱

داده‌های بالا را می‌توان به صورت زیر رتبه‌بندی نمود.

تیمارها و رتبه‌ها							
تیمار ۱	رتبه	تیمار ۲	رتبه	تیمار ۳	رتبه	تیمار ٤	رتبه
۲۲/۲	۷	۲٤/۱	۱۱	۲۵/۹	۱٤	۲۳/۹	۱۰
۱۷/۳	۲	۳۰/۳	۱۹	۱۸/٤	۳	۲٤/۸	۱۲
۲۱/۲	٤	۲۷/٤	۱۷	۲۳/۲	۹	۲۸/۲	۱۸
۲۵/۲	۱۳	۲٦/٤	۱۵/۵	۲۱/۹	٦	۲٦/٤	۱۵/۵
۲٦/٤	۱	۳٤/۸	۲۰	۲۲/٦	۸	۲۱/۷	۵
R_i	۲۷		۸۲/۵		٤۰		٦۰/۵
R_i	۵/٤		۱٦/۵		۸		۱۲/۱

آماره آزمون بصورت زیر محاسبه می‌شود.

$$\begin{aligned} H &= \frac{۱۲}{N(N+۱)}\sum\frac{R_i^{۲}}{n_i} - ۳(N+۱) \\ &= \frac{۱۲}{۲۰(۲۱)}\left(\frac{۲۷^{۲}}{۵} + \frac{۸۲/۵^{۲}}{۵} + \frac{٤۰^{۲}}{۵} + \frac{٦۰/۵^{۲}}{۵}\right) - ۳\times ۲۱ \\ &= \frac{۱۲}{٤۲۰}(۲۵۵۹/۱) - ۷۸ = ۷۳/۱۱۷ - ٦۳ = ۱۰/۱۲ \end{aligned}$$

با توجه به اینکه آماره H محاسبه شده از $\chi^{۲}_{(٤,\ ۰/۰۵)} = ۹/٤۹$ بزرگتر است، فرض صفر مبنی بر برابر بودن میانه‌های جوامع رد می‌شود.

مقدار احتمال که در خروجی برنامه Minitab (جدول ۱۲-٦) مشاهده می‌شود (۰/۰۱۸)، با مقدار احتمال به دست آمده در تحلیل واریانس یک‌طرفه قابل مقایسه خواهد بود. با توجه به اینکه فقط یک رتبه گره‌خورده در داده‌ها وجود دارد (۱۵/۵)، مقدار آماره H با مقدار تصحیح شده آن تفاوت چندانی ندارد. مقدار آماره z در خروجی نشان می‌دهد که میانگین رتبه‌ها در هر گروه چقدر با میانگین کل رتبه‌ها تفاوت دارد. z مربوطه در مورد تیمار اول به صورت

$$z_۱ = \frac{\overline{R}_i - \overline{R}}{\sqrt{(N+۱)(N/n_i - ۱)/۱۲}} = \frac{۵/٤ - ۱۰/٤}{\sqrt{۲۱ \times ۳ / ۱۲}} = \frac{-۵/۱}{۲/۲۹۱} = -۲/۲۳$$

به دست می‌آید. این عدد نشان می‌دهد که میانگین رتبه‌های تیمار اول به طور قابل ملاحظه‌ای از میانگین کل رتبه‌ها کمتر است.

جدول ۱۲-٦. حل مثال ۱۲-٦ و خروجی مربوطه در برنامه Minitab.

مراحل:

۱. داده‌های نمونه‌ها را در زیر هم در ستون دوم وارد کرده و در مقابل هر داده در ستون اول شماره نمونه را وارد کنید.

۲. وارد مسیر زیر شوید.

Stat > Nonparametrics > Kruskal-Wallis...

۳. در مقابل گزینه Response ستون دوم و در مقابل Factor ستون اول را وارد کنید.

۴. گزینه OK را انتخاب کنید.

```
Kruskal-Wallis Test on C2

C1         N  Median  Ave Rank      Z
1          5   21.20       5.4  -2.23
2          5   27.40      16.5   2.62
3          5   22.60       8.0  -1.09
4          5   24.80      12.1   0.70
Overall   20              10.5

H = 10.12  DF = 3  P = 0.018
H = 10.12  DF = 3  P = 0.018  (adjusted for ties)
```

۱۲-٦ آزمون فریدمن (Friedman)

این روش یک جایگزین ناپارامتری برای تحلیل طرح بلوک‌های کامل تصادفی است. به ویژه هنگامی که فرض‌های نرمال بودن و یکنواختی واریانس‌ها برقرار نباشد. فرض صفر در روش فریدمن بیان می‌کند که جوامع مربوط به تیمارها با هم متفاوت نیستند. اگر فقط دو تیمار وجود داشته باشد، از آزمون نمونه‌های جفت شده ویلکاکسون برای مقایسه استفاده می‌شود. برای انجام این آزمون فرض‌هایی باید برقرار باشند از جمله اینکه ۱- تیمارها درون هر بلوک به طور تصادفی به واحدهای آزمایشی منتسب شده باشند. ۲- داده‌های درون هر بلوک را بتوان رتبه‌بندی نمود. ۳- توزیع احتمال جوامعی که هر بلوک از آن استخراج می‌شود پیوسته باشد.

در اینجا تعداد تیمارها با k و تعداد بلوک‌ها با n نشان داده می‌شود. داده‌های درون هر بلوک از ۱ تا k رتبه‌گذاری می‌شوند سپس جمع رتبه‌ها برای هر تیمار جداگانه محاسبه و با R_i (برای تیمار iام) نشان داده می‌شود. نهایتاً آماره آزمون به صورت

$$S = \frac{۱۲}{nk(k+۱)} \sum [R_i - n(k+۱)\,/\,۲]^۲ = \frac{۱۲}{nk(k+۱)} \sum {R^۲}_i - ۳n(k+۱)$$

محاسبه می‌شود. تحت فرض صفر اگر n یا k از ۵ کمتر نباشند، آماره S تقریباً دارای توزیع کای‌دو با $k-۱$ درجه آزادی است. آزمون فریدمن نیز آزمونی یک طرفه بوده و فرض صفر در آن هنگامی رد می‌شود که مقدار آماره S مساوی یا بزرگتر از $\chi^۲_{(k-۱,\ \alpha)}$ باشد.

مثال ۱۲-۷. نتایج مقایسه عملکرد چهار واریته گندم بر حسب تن در هکتار که در قالب طرح بلوک‌های کامل تصادفی پیاده شده، بصورت زیر بوده است.

	واریته ۱	واریته ۲	واریته ۳	واریته ٤
بلوک ۱	(۲) ۷/٤	(٤) ۹/۸	(۱) ۷/۳	(۳) ۹/٥
بلوک ۲	(۲) ٦/٥	(۳) ٦/۸	(۱) ٦/۱	(٤) ۸/۰
بلوک ۳	(۱) ٥/٦	(۲) ٦/۲	(۳) ٦/٤	(٤) ۷/٤
جمع رتبه	٥	۹	٥	۱۱
میانگین رتبه	۱/٦۷	۳	۱/٦۷	۳/٦۷

آماره S را می‌توان به صورت

$$S = \frac{۱۲}{۳\times ٥\times ٤}[٥^۲ + ۹^۲ + ٥^۲ + ۱۱^۲] - ۳\times ۳\times ٥ = \frac{۱۲}{٦۰}\times ۲٥۲ - ٤٥ = ٥۰/٤ - ٤٥ = ٥/٤$$

محاسبه نمود. با توجه به اینکه مقدار آماره S محاسبه شده از $\chi^۲_{(۳,\ ۰/۰٥)} = ۷/۸۱٥$ کوچکتر است، پس دلیل کافی برای رد فرض صفر (برابر بودن میانه‌های جوامع) وجود ندارد. مقدار احتمال در خروجی Minitab (۰/۱٤٥) با مقدار احتمال به دست آمده در روش تحلیل واریانس طرح بلوک‌های کامل تصادفی قابل مقایسه خواهد بود.

جدول ۱۲-۷. حل مثال ۱۲-۷ و خروجی مربوطه در برنامه Minitab.

مراحل:

۱. داده‌ها را در ستون سوم وارد کرده و در مقابل هر داده در ستون اول شماره تیمار و در ستون دوم شماره بلوک مربوطه را وارد کنید.

۲. وارد مسیر زیر شوید.

Stat > Nonparametrics > Friedman...

۳. در مقابل گزینه Response ستون سوم، در مقابل Treatment ستون اول و در مقابل Blocks ستون دوم را انتخاب کنید.

۴. گزینه OK را انتخاب کنید.

Friedman Test: Response versus Treatment blocked by Block

```
S = 5.40  DF = 3  P = 0.145

                     Est   Sum of
Treatment  N  Median   Ranks
1          3   6.200     5.0
2          3   6.800     9.0
3          3   6.100     5.0
4          3   8.000    11.0

Grand median = 6.775
```

برنامه Minitab مقدار آماره S و مقدار احتمال را متناسب با رتبه‌های گره‌خورده تصحیح می‌کند. البته در این مثال رتبه گره‌خورده وجود ندارد. درجه آزادی برای آماره S نیز برابر با تعداد تیمارها منهای یک است. همچنین برنامه Minitab میانه را برای هر تیمار (۶/۷۷۵ به اضافه اثر تیمار) محاسبه می‌کند. N نیز در خروجی برنامه تعداد بلوک‌ها را نشان می‌دهد.

۱۲-۷ ضریب همبستگی رتبه‌ای اسپیرمن

ضریب همبستگی رتبه‌ای اسپیرمن (r_S) شاخصی است که ارتباط خطی بین رتبه‌ها را نشان می‌دهد. ضریب همبستگی رتبه اسپیرمن همانند ضریب همبستگی پیرسون محاسبه می‌شود با این تفاوت که در آن از رتبه‌ها استفاده می‌شود.

مثال ۱۲-۸ فرض کنید ۱۰ تابلوی نقاشی توسط دو متخصص رتبه‌گذاری شده و نتایج زیر بدست آمده است.

شماره تابلو	امتیاز متخصص ۱	امتیاز متخصص ۲	تفاوت بین رتبه‌ها (d)	d^2
۱	٤	٥	−۱	۱
۲	۱	۲	−۱	۱
۳	۹	۱۰	−۱	۱
٤	٥	٦	−۱	۱
٥	۲	۱	۱	۱
٦	۱۰	۹	۱	۱
۷	۷	۷	۰	۰
۸	۳	۳	۰	۰
۹	٦	٤	۲	٤
۱۰	۸	۸	۰	۰
				$\sum d^۲ = ۱۰$

ضریب همبستگی رتبه اسپیرمن را در مورد مثال اخیر می‌توان به صورت زیر محاسبه نمود.

$$r_s = \frac{SS_{uv}}{\sqrt{SS_{uu}SS_{vv}}} = \frac{۷۷/٥}{\sqrt{۸۲/٥ \times ۸۲/٥}} = ۰/۹۳۹$$

ضریب همبستگی رتبه اسپیرمن را از طریق فرمول زیر نیز بدست می‌آید.

$$r_s = ۱ - \frac{٦\sum d^۲}{n(n^۲-۱)} = ۱ - \frac{٦(۱۰)}{۱۰(۹۹)} = ۱ - \frac{٦}{۹۹} = ۰/۹۳۹$$

جدول ۱۲–۸ حل مثال ۱۲–۸ و خروجی مربوطه در برنامه Minitab.

مراحل:

۱. داده‌های نمونه‌های اول و دوم را بترتیب را در ستون‌های اول و دوم وارد کنید.

۲. در برنامه Minitab فقط ضریب همبستگی پیرسون محاسبه می‌شود لذا اگر داده‌های خام به صورت رتبه نیستند، با استفاده از دستور Rank در مسیر Calc > Caculator داده‌های هر ستون را رتبه‌بندی کنید.

۳. وارد مسیر زیر شده و متغیرها (ستون‌های حاوی رتبه‌ها) را انتخاب کنید:

Stat > Basic Statistics > Correlation...

۴. گزینه OK را انتخاب کنید.

```
Pearson correlation of C1 and C2 = 0.939
P-Value = 0.000
```

با توجه به اینکه برنامه Minitab دستور جداگانه‌ای برای محاسبه ضریب همبستگی اسپیرمن ندارد، مقدار احتمال به دست آمده در این حالت قابل تفسیر نخواهد بود.

ضریب همبستگی رتبه‌ای اسپیرمن

$$r_s = \frac{SS_{uv}}{\sqrt{SS_{uu}SS_{vv}}}$$

که در آن:

$$SS_{uv} = \sum(u_i - \overline{u})(v_i - \overline{v}) = \sum u_i v_i - \frac{(\sum u_i)(\sum v_i)}{n}$$

$$SS_{uu} = \sum(u_i - \overline{u})^۲ = \sum u_i^۲ - \frac{(\sum u_i)^۲}{n}$$

$$SS_{vv} = \sum(v_i - \overline{v})^۲ = \sum v_i^۲ - \frac{(\sum v_i)^۲}{n}$$

u_i رتبه iامین داده در نمونه اول و v_i رتبه iامین داده در نمونه دوم است. n نیز تعداد جفت مشاهدات و یا تعداد مشاهدات در هر نمونه است.

البته ضریب همبستگی رتبه را به صورت زیر نیز می‌توان به دست آورد:

$$r_s = ۱ - \frac{۶\sum d^۲}{n(n^۲ - ۱)}$$

این فرمول در صورت وجود داده‌های گره خورده کاملاً دقیق نیست. منتها در صورتی که نسبت داده‌های گره‌خورده کم باشد تقریب خوبی به دست خواهد داد.

در محاسبه ضریب همبستگی رتبه‌ای اسپیرمن فرض بر این است که نمونه‌ها به صورت تصادفی و مستقل از هم از جامعه استخراج شده‌اند و توزیع احتمال جوامعی که نمونه‌ها از آن استخراج شده‌اند پیوسته می‌باشد.

مثال ۱۲-۹. تعداد سیگارهای مصرف شده توسط ۱۵ زن سیگاری باردار در دوران بارداری و وزن نوزادان آنها در بدو تولد به صورت زیر بوده است. ضریب همبستگی اسپیرمن را محاسبه و تفسیر نمایید.

	تعداد سیگار در روز	رتبه	وزن نوزاد	رتبه	d	d^2
۱	۱۲	۱	۳/۲	۵	$-$٤	۱٦
۲	۱۵	۲	۳/٦	۹	$-$۷	٤۹
۳	۳۵	۱۳	۳/۰	٤	۹	۸۱
٤	۲۱	۷	۳/۷	۱۰	$-$۳	۹
۵	۲۰	۵/۵	۳/۹	۱۳/۵	$-$۸	٦٤
٦	۱۷	۳	۳/۸	۱۱/۵	$-$۸/۵	۷۲/۲۵
۷	۱۹	٤	٤/۲	۱۵	$-$۱۱	۱۲۱
۸	٤٦	۱۵	۳/۳	٦	۹	۸۱
۹	۲۰	۵/۵	۳/۸	۱۱/۵	$-$٦	۳٦
۱۰	۲۵	۸/۵	۲/۳	۱	۷/۵	۵٦/۲۵
۱۱	۳۹	۱٤	۲/۹	۳	۱۱	۱۲۱
۱۲	۲۵	۸/۵	۳/٤	۷	۱/۵	۲/۲۵
۱۳	۳۰	۱۲	۳/۵	۸	٤	۱٦
۱٤	۲۷	۱۰	۲/۸	۲	۸	٦٤
۱۵	۲۹	۱۱	۳/۹	۱۳/۵	$-$۲/۵	٦/۲۵
						جمع = ۷۹۵

ضریب همبستگی اسپیرمن به صورت زیر محاسبه می‌شود.

$$r_s = ۱ - \frac{٦\sum d^۲}{n(n^۲-۱)} = ۱ - \frac{٦(۷۹۵)}{۱۵(۱۵^۲-۱)} = ۱ - ۱/٤۲ = -۰/٤۲$$

منفی بودن ضریب همبستگی اسپیرمن نشان می‌دهد که با افزایش تعداد سیگارهای مصرف شده توسط مادران باردار وزن نوزادان کاهش می‌یابد. ضریب همبستگی اسپیرمن محاسبه شده از طریق نمونه (r_s) می‌تواند برای آزمون فرض در مورد ضریب همبستگی اسپیرمن جامعه (ρ_s) به کار رود. در مثال اخیر فرض‌ها را می‌توان به صورت

$$H_\circ : \rho_s = ۰$$
$$H_۱ : \rho_s < ۰$$

نوشت. جدول ۱۲–۹ مقادیر بحرانی ضریب همبستگی اسپیرمن را با توجه به اندازه نمونه نشان می‌دهد. مشاهده می‌شود که برای $n = ۱۵$ و $\alpha = ۰/۰۵$، عدد بحرانی برابر با ۰/٤٤۱ است. به عبارت دیگر درصورت صحت فرض صفر $(\rho_s = ۰)$، احتمال مشاهده r_s بزرگتر از ۰/٤٤۱ در یک نمونه ۱۵ عضوی برابر با ۰/۰۵ است. به همین ترتیب احتمال مشاهده r_s کوچکتر از ۰/٤٤۱– نیز برابر با ۰/۰۵ است. با توجه به اینکه در نمونه اخیر r_s برابر با ۰/٤۲– است لذا دلیل کافی برای رد فرض صفر وجود نداشته و بر مبنای نمونه موجود بین تعداد سیگارهای مصرف شده و وزن نوزاد در بدو تولد ارتباط معکوس معنی‌دار وجود ندارد. شاید استفاده از یک نمونه

بزرگتر و وارد نمودن عوامل دیگری از جمله وزن پدر و مادر و جنسیت نوزاد ارتباط مورد بررسی را روشن سازد.

جدول ۱۲-۹. مقادیر بحرانی ضریب همبستگی اسپیرمن در آزمون یک‌طرفه برای آزمون فرض $H_\circ : \rho_s = ۰$.

	سطح α			
n	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵
۵	۰/۹۰۰	–	–	–
۶	۰/۸۲۹	۰/۸۸۶	۰/۹۴۳	–
۷	۰/۷۱۴	۰/۷۸۶	۰/۸۹۳	–
۸	۰/۶۴۳	۰/۷۳۸	۰/۸۳۳	۰/۸۸۱
۹	۰/۶۰۰	۰/۶۸۳	۰/۷۸۳	۰/۸۳۳
۱۰	۰/۵۶۴	۰/۶۴۸	۰/۷۴۵	۰/۷۹۴
۱۱	۰/۵۲۳	۰/۶۲۳	۰/۷۳۶	۰/۸۱۸
۱۲	۰/۴۹۷	۰/۵۹۱	۰/۷۰۳	۰/۷۸۰
۱۳	۰/۴۷۵	۰/۵۶۶	۰/۶۷۳	۰/۷۴۵
۱۴	۰/۴۵۷	۰/۵۴۵	۰/۶۴۶	۰/۷۱۶
۱۵	۰/۴۴۱	۰/۵۲۵	۰/۶۲۳	۰/۶۸۹
۱۶	۰/۴۲۵	۰/۵۰۷	۰/۶۰۱	۰/۶۶۶
۱۷	۰/۴۱۲	۰/۴۹۰	۰/۵۸۲	۰/۶۴۵

	سطح α			
n	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵
۱۸	۰/۳۹۹	۰/۴۷۶	۰/۵۶۴	۰/۶۲۵
۱۹	۰/۳۸۸	۰/۴۶۲	۰/۵۴۹	۰/۶۰۸
۲۰	۰/۳۷۷	۰/۴۵۰	۰/۵۳۴	۰/۵۹۱
۲۱	۰/۳۶۸	۰/۴۳۸	۰/۵۲۱	۰/۵۷۶
۲۲	۰/۳۵۹	۰/۴۲۸	۰/۵۰۸	۰/۵۶۲
۲۳	۰/۳۵۱	۰/۴۱۸	۰/۴۹۶	۰/۵۴۹
۲۴	۰/۳۴۳	۰/۴۰۹	۰/۴۸۵	۰/۵۳۷
۲۵	۰/۳۳۶	۰/۴۰۰	۰/۴۷۵	۰/۵۲۶
۲۶	۰/۳۲۹	۰/۳۹۲	۰/۴۶۵	۰/۵۱۵
۲۷	۰/۳۲۳	۰/۳۸۵	۰/۴۵۶	۰/۵۰۵
۲۸	۰/۳۱۷	۰/۳۷۷	۰/۴۴۸	۰/۴۹۶
۲۹	۰/۳۱۱	۰/۳۷۰	۰/۴۴۰	۰/۴۸۷
۳۰	۰/۳۰۵	۰/۳۶۴	۰/۴۳۲	۰/۴۷۸

تمرین‌ها

۱۲-۱. در چه شرایطی استفاده از آزمون علامت به جای آزمون t ترجیح داده می‌شود؟ شرط استفاده از آزمون علامت چیست؟

۱۲-۲. در یک نظرسنجی از ۱۸ نفر برای بررسی مزه دو نوع بیسکویت، ۱۳ نفر نوع A و ۶ نفر نوع B را ترجیح دادند. با استفاده از آزمون علامت، مقدار احتمال را در آزمون فرض H_1 مبنی بر برتری نوع A به دست آورید.

۱۲-۳. در مصاحبه‌ای با ۲۴ زوج، زن و مرد به طور مستقل به این سؤال که مایلید چند بچه داشته باشید جواب داده‌اند و نتایج زیر به دست آمده است. **الف)** آیا عقیده زنان و مردان در مورد تعداد فرزندان متفاوت است؟ آزمون علامت را در سطح $\alpha = ۰/۰۵$ انجام دهید. **ب)** از آزمون رتبه علامت‌دار ویلکاکسون برای آزمون فرض مذکور استفاده نمایید.

زوج	جواب زن	جواب مرد
۱	۲	۳
۲	۱	۱
۳	۱	۲
٤	۳	۲
٥	۱	٥
٦	۲	۰
۷	۳	۰
۸	۲	۱
۹	۱	۳
۱۰	۲	۳
۱۱	۲	٤
۱۲	۳	۱

زوج	جواب زن	جواب مرد
۱۳	۱	۲
۱٤	۲	۳
۱٥	۲	۲
۱٦	۰	۰
۱۷	۲	۱
۱۸	۱	۲
۱۹	۲	۳
۲۰	۳	٤
۲۱	۱	۳
۲۲	۰	۰
۲۳	۲	۱
۲٤	۱	۱

۱۲–۴. رتبه‌های داده شده بر مبنای یک مصاحبه و نمرات استعدادسنجی به ٦ متقاضی استخدام به صورت زیر بوده است. ضریب همبستگی رتبه‌ای اسپیرمن را به دست آورده و تفسیر کنید.

رتبه مصاحبه	٤	٦	۱	۳	۲	٥
نمره استعداد	۳۸	٥٦	۲۸	۲۹	۳۲	٤۷

۱۲–۵. آفلاتوکسین ماده‌ای شدیداً سمی است که در هوای گرم و آفتابی در مزارع ذرت توسط نوعی کپک بر روی بلال تولید شده و برای کشاورزان خطرناک است. اثر سه نوع اسپری A، B و C بر کنترل آفلاتوکسین بررسی گردید بدین ترتیب که ۱۰ بوته به صورت تصادفی از یک مزرعه آلوده برداشت و هر کدام به سه قسمت تقسیم گردید. هر قسمت نیز در قالب آزمایش بلوک‌های کامل تصادفی تحت تأثیر یک نوع اسپری قرار گرفت. میزان آفلاتوکسین (برحسب قسمت در بیلیون یا PPB) پس از مدت زمان معین از هر قسمت خوشه اندازه‌گیری شد که نتایج آن در زیر آمده است.

خوشه	اسپری A	اسپری B	اسپری C
۱	۲۱	۲۳	۱٥
۲	۲۹	۳۰	۲۱
۳	۱٦	۱۹	۱۸
٤	۲۰	۱۹	۱۸
٥	۱۳	۱۰	۱٤
٦	٥	۱۲	٦
۷	۱۸	۱۸	۱۲
۸	۲٦	۳۲	۲۱
۹	۱۷	۲۰	۹
۱۰	٤	۱۰	۲

الف) آیا دلایل کافی در سطح $\alpha = ۰/۰۵$ مبنی بر متفاوت بودن اثر اسپری‌ها در کنترل آفلاتوکسین وجـود دارد؟ **ب)** در صورت مثبت بودن جواب قسمت الف، اسپری‌ها را دوتا دوتا با یک روش ناپارامتری مناسـب در سـطح $\alpha = ۰/۰۱$ مقایسه کرده و بهترین اسپری را در کنترل آفلاتوکسین معرفی کنید. ج) چه فرض‌هایی بـرای معتبـر بودن نتایج آزمون‌های بالا باید برقرار باشد؟

۶-۱۲. میزان عنصر سدیم در برگ‌های سه رقم گندم بر حسب میکرومول در هر گرم برگ خشک به صـورت زیر بوده است. از هر رقم چهار نمونه بررسی گردید. آیا در سطح ۵٪ می‌توان گفت که محتوای سدیم برگ بـین ارقام متفاوت بوده است؟

	میزان سدیم	
رقم اول	رقم دوم	رقم سوم
۸۸/۹۹	۸٤/۰٤	۸٥/۲٥
۸٥/۱٥	۸٥/۱۳	۸٤/٤۳
۸٤/۷۲	۸٤/٤۸	۸۳/۷۰
۸٤/۲۰	۸٤/۱۰	۸٦/۰۹