بسم الله الرحمن الرحیم

نمی دانم که در طرح بزرگ خدا من چه نقشی دارم و چه سرنوشتی؟ ولی

این قدر مطمئنم که بی هیچ نیست.

الهی ای خداوند شب و روز   ز شوق بندگی جانم بیفروز

خداوندا برافروزان روانم   بزن از عشق خود آتش به جانم

تقدیم به

پدر و مادر بزرگوارم

آنان که در مقابل سختی‌ها و ناملایمات روزگار ماندند و از عدسی کوچک دوربین زندگی، هزاران پنجره‌ی تماشا و کهکشانی از ستاره‌های امید و پویایی را به رویم گشودند و با واژه‌ای به بلندای همت و صمیمیت، خویشتن را برای آرامش و سربلندیم صرف کردند؛ دستشان را می‌بوسم.

خواهران گلم: فاطمه، خدیجه و لیلا

ستاره‌های پرفروغ زندگیم؛ آنان که وجودشان برای من مایه‌ی آرامش و مهربانی است.

الهی

وای بر من، اگر دانشم رهزنم شود و کتابم حجابم

اکنون که پایان نامه ام را با موفقیت گذرانده ام، با تمام وجود از الهه ی دانش و دانایی، خدای مهربانم، شکرگزاری می کنم و بر خود لازم می دانم از استاد دوست داشتنی ام سرکار خانم دکتر نرگس عباسی، که با تعهد و عشق فراوان، نه تنها در مقام استاد راهنمای پایان نامه ام، بلکه در طول دوره ی کارشناسی ارشد با ارشادهای ارزنده ی خویش اینجانب را در مسیر صحیح دانش آموختگی ام هدایت کردند، بی نهایت قدردانم و بهترین ها را برایشان آرزومندم.

از جناب آقای دکتر مسعود یارمحمدی، که با قبول مشاوره ی پایان نامه ام و با در اختیار قرار دادن دانششان، مرا مرهون لطف خویش قرار دادند، بی نهایت سپاسگزارم و تندرستی شان را از درگاه الهی خواهانم.

از جناب آقای دکتر علی شادرخ و دکتر امیر تیمور پاینده، از بزرگان آمار ایران، قدردان و متشکرم و برای من افتخاری بود که ایشان داوری پایان نامه ام را قبول کردند.

از ناظر تحصیلات تکمیلی جناب آقای دکتر پرویز نصیری قدردان و متشکرم که نظارت جلسه ی دفاع بنده را متقبل شدند.

از کلیه ی مدرسان دوران ۱۲ ساله ی تحصیلم، کلیه ی اساتید دانشگاه پیام نور مرکز صفاشهر، به ویژه دکتر احد ملک زاده و از اساتید این مرکز و مرکز شیراز، به ویژه دکتر نرگس عباسی که از محضرشان علم آموختم؛ بی نهایت سپاسگزار و قدردانم.

و در پایان از دوستان خوبم: فاطمه حبیبی، صدیقه همتی، زهرا راس روش، طاهره صابر، افسانه کرمجانی، فاطمه قوتی، کامران قادری، جواد شعبانی و ابراهیم اسلوب که همراهان واقعی دوران دانشجویی بودند، بسیار ممنونم و سلامت و سرافرازیشان را از خداوند منان خواستارم.

چکیده

در تحلیل واریانس کلاسیک، پراکندگی با در نظر گرفتن مربع فاصله مشاهدات نمونه از میانگین نمونه اندازه‌گیری می‌شود. همچنین می‌توان از حل معادله‌ی درجه دوم تفاضل مشاهدات از یک ماتریس ثابت (که بعد آن به اندازه‌ی نمونه بستگی دارد)، فرمول دیگری برای واریانس و انجام مطالعات بیشتر ارائه داد. استفاده از میانگین در محاسبه‌ی واریانس افزایش دقت در محاسبه را به‌دنبال دارد، به‌خصوص اگر برنامه‌های رایانه‌ایی به‌کار گرفته شود. یک روش پیشنهادی دیگر برای محاسبه‌ی واریانس، استفاده از توزیع فراوانی، مورد بحث قرار می‌گیرد در این تحقیق معیاری از اختلاف مرتبه‌ی اول مشاهدات و استفاده از توزیع فراوانی در محاسبه‌ی پراکندگی ارائه می‌شود. همچنین معیاری از پراکندگی برای متغیر پاسخ یک‌متغیره یا چندمتغیره براساس تمام جفت فواصل بین عناصر نمونه در نظر می‌گیریم و برای تجزیه‌ی مؤلفه‌ی فواصل مشابه، توان‌های در بازه‌ی [۰,۲) را نتیجه می‌گیریم. در تحلیل آنالیز واریانس، آماره‌ی F هنگامی‌که توان دو است، به‌دست می‌آید. برای هر مقدار در بازه‌ی (۰,۲) این تجزیه یک آزمون ناپارامتری برای فرض برابری توزیع در نمونه‌های چندتایی تعیین می‌کند.

فهرست

فصل اول

واریانس نمونه‌ای و اختلاف مرتبه‌ی اول مشاهدات

مقدمه

از آن‌جایی که روش معمول محاسبه‌ی واریانس نمونه وابسته به اختلاف حول میانگین نمونه است، ممکن است فاقد دقت و صراحت باشد، به‌ویژه هنگامی که برنامه‌های رایانه‌ای برای محاسبه مورد استفاده قرار می‌گیرد. مشکل محاسبه‌ی واریانس زمانی که از میانگین استفاده می‌کنیم، توسط راس[1] در سال ۱۹۸۷ مطرح شد. برخی راه حل‌های این مشکل توسط جواردر[2] در سال ۲۰۰۲ مورد بحث قرار گرفت.

واریانس $S_n^۲$، n مشاهده‌ی $a_۱, \dots, a_n$، از مجموع مربعات انحراف مشاهدات از میانگین (TSS) تقسیم بر درجه آزادی به‌دست می‌آید، که در آن

$$\begin{aligned} TSS &= (n-۱)S_n^۲ \\ &= \sum_{i=۱}^{n}(a_i-\bar{a})^۲ = \sum_{i=۱}^{n} a_i^۲ - n\bar{a}^۲, \quad n \geq ۲ \end{aligned} \tag{۱}$$

اگر تعداد مشاهدات نمونه کوچک باشد، آخرین عبارت دررابطه‌ی (۱) را می‌توان به‌طور ذهنی انجام داد.

مقدار TSS را همچنین می‌توان به شکل عبارت زیر بیان کرد:

$$\begin{aligned} TSS &= \frac{۱}{n}\sum_{۱\leq i<j\leq n}(a_i-a_j)^۲ \\ &= \frac{۱}{۲n}\sum_{i=۱}^{n}\sum_{j=۱}^{n}(a_i-a_j)^۲ = \frac{۱}{n}\sum_{i=۱}^{n}\sum_{j=۱}^{i-۱}(a_i-a_j)^۲ \end{aligned} \tag{۲}$$

مفهوم رابطه‌ی(۲) این است که واریانس نمونه ایی n مشاهده را می‌توان به‌آسانی با محاسبه‌ی پراکندگی $\binom{n}{۲}$ جفت مشاهده و سپس محاسبه‌ی میانگین آن‌ها بدست آورد، به‌طوری‌که برای واریانس نمونه‌ی به حجم $n \geq ۲$ داریم:

$$s_n^۲ = \binom{n}{۲}^{-۱}\sum_{i=۱}^{n}\sum_{j=۱}^{i-۱}\frac{w_{ij}^۲}{۲} \tag{۳}$$

[1] Ross

[2] Joatder

که در آن $w_{ij} = a_i - a_j;\ i,j = ۱, \dots, n$. در این عبارت برای محاسبه‌ی واریانس، از اختلاف مرتبه‌ی اول مشاهدات استفاده شده است. همچنین واریانس می‌توان توسط توان دوم تفاوت مشاهدات نمونه از ماتریس ثابت C (که بعد آن به اندازه‌ی نمونه بستگی دارد) بیان شود که در بحث‌های بعدی مطرح می‌شود. یک روش پیشنهادی دیگر برای محاسبه‌ی واریانس از طریق توزیع فراوانی است.

در تحلیل واریانس کلاسیک یک‌متغیره و تحلیل واریانس چندمتغیره اگر $\mu_۱, \dots, \mu_K$ میانگین‌ها یا بردار میانگین جامعه باشد، فرض برابری میانگین جامعه به‌صورت

$$H_۰: \mu_۱ = \cdots = \mu_K \tag{۴}$$

در برابر فرض مقابل $\mu_j \neq \mu_K$ به ازای حداقل یک $j \neq k$ است. این استنباط نیازمند این است که خطای تصادفی دارای توزیع نرمال با میانگین صفر و واریانس ثابت باشد (به ککران[۳] و کاکس[۴] (۱۹۵۷)، سرل[۵]، کسلا[۶] و مک‌کولوچ[۷] (۱۹۹۲)، هند[۸] و تیلور[۹] (۱۹۸۷)، ماردیا[۱۰]، کنت[۱۱] و بای بی[۱۲] (۱۹۷۹) مراجعه شود).

در تحلیل واریانس، پراکندگی کل مشاهدات متغیر پاسخ به مجموع مربعات تیمارها (TSS) و مجموع مربعات خطا بین نمونه‌ایی (SSE) تقسیم می‌شود. هنگامی‌که فرض رایج نرمال بودن

---

[۳] Cochran

[۴] Cox

[۵] Searle

[۶] Casella

[۷] McCulloch

[۸] Hand

[۹] Taylor

[۱۰] Mardia

[۱۱] Kent

[۱۲] Bibby

توزیع خطا وثابت بودن واریانس آن برقرار باشد تحت فرض صفر توزیع‌ها یکسانند، و تحت فرض مقابل توزیع‌ها تنها در پارامتر مکانی متفاوتند (بعد از تبدیل یکسان هستند).

اگر توزیع‌ها تنها در پارامتر مکانی متفاوت باشند، برای یک متغیر پاسخ از روش‌هایی براساس آزمون‌های ناپارامتری رتبه‌ی کروسکال – والیس[۱۳] و یا آزمون‌های میانه‌ی مود[۱۴] می‌توان برای آزمون فرض برابری میانه جامعه به‌کار برد.

در مواردی که فرض نرمال بودن یا ثابت بودن واریانس خطا برداشته شود می‌توان از آماره‌ی F توسط یک روش آزمون تبدیل استفاده کرد (افرون[۱۵] و تیب شراین[۱۶] (۱۹۹۳) فصل ۱۵، داویسون[۱۷] و هینکلی[۱۸] فصل ۴ (۱۹۹۷)).

اگرچه در توزیع‌های تجربی با میانگین‌های برابر ممکن است ویژگی‌ها از یکدیگر متفاوت باشند، برای فرض (۴)، آماره‌ی F به‌کار می‌رود. تحلیل واریانس یک‌متغیره و چندمتغیره را برای آزمون فرض کلی (۵) با کمک گرفتن از تجربه برای دیگر توان‌های فواصل تعمیم می‌دهیم.

برای K نمونه‌ی تصادفی مستقل، با تابع توزیع تجمعی به‌ترتیب $F_۱, \dots, F_K$ فرض برابری توزیع‌ها به‌صورت

$$H_۰: F_۱ = \cdots = F_K \qquad (۵)$$

در برابر فرض مقابل $F_j \neq F_k$ حداقل به ازای یک $۱ \leq j < k \leq K$، در نظر گرفته می‌شود، که در آن فرض می‌کنیم هر K متغیر تصادفی از مقادیر R^p برای $p \geq ۱$ آمده‌اند و توزیع F_jها نامشخص‌اند.

[۱۳] Kruskal-Wallis

[۱۴] Mood

[۱۵] Efron

[۱۶] Tibshirani

[۱۷] Dvison

[۱۸] Hinkley

در این پایان‌نامه یک روش جدید (مؤلفه‌ی فاصله) برای اندازه‌گیری پراکندگی کل نمونه‌ها پیشنهاد می‌شود، که اجازه دارد در آن یک تقسیم‌بندی از پراکندگی کل برای مؤلفه‌های واریانس طبق جدول تحلیل آنالیز واریانس انجام دهد. در ادامه مؤلفه‌ی فاصله یک آماره‌ی آزمون برای فرض کلی (۵) برابری توزیع‌ها تعیین می‌کند، و همچنین معیاری از پراکندگی براساس یک فاصله‌ی اقلیدسی بین جفت عناصر نمونه که برای هر توان $\alpha \epsilon (\circ, ۲]$ است، ارائه می‌دهد.

در حالت خاصی که $\alpha = ۲$، مجموع مربعات خطا در جدول تحلیل واریانس به‌دست می‌آید. برای تمام مقادیر α در فاصله‌ی $(\circ,۲)$ تجزیه‌ای به‌دست می‌آورد به‌طوری‌که بر طبق آماره‌ی F، یک آماره‌ی آزمون برای فرض کلی (۵) تعیین می‌شود.

اکریتاس[۱۹] و آرنولد[۲۰] (۱۹۹۴) یک مدل کلی برای ساختار داده‌هایی که توزیع‌شان براساس توزیع متغیر پاسخ مدل‌بندی شده‌اند را پیشنهاد دادند. فرض این که اثر تیماری یا اثر متقابل بر طبق توزیع در مدل نباشد، برابر صفر است. برای مثال به برونر[۲۱] و پوری[۲۲] (۱۹۹۴) مراجعه کنید.

فاصله‌ی دیگر براساس تقریب به آزمون (۴) یا (۵) در ویرایش آخر گوئر[۲۳] و کرزانوسکی[۲۴] (۱۹۹۹) و اندرسون[۲۵] (۲۰۰۱) معرفی شده‌اند که در بوم‌شناسی، اقتصاد، ژنتیک (اسکوفیر[۲۶]،

[۱۹] Akritas

[۲۰] Arnold

[۲۱] Brunner

[۲۲] Puri

[۲۳] Gower

[۲۴] Krzanowski

[۲۵] Anderson

[۲۶] Excoffier

اسموز[۲۷]، کواترو[۲۸] (۱۹۹۲)، مک‌آردل[۲۹] و اندرسن (۲۰۰۱)، زپلا[۳۰] و اسچورک[۳۱] (۲۰۰۶)). به‌کار می‌روند.این روش‌ها از شیوه‌ی پیشنهاد داده شده متفاوت است، اختلاف در این است که آن‌ها مربع فاصله را با یک روش دیگری از تجزیه‌ی فاصله‌ها، یا یک معیار نامتشابه از توان فاصله‌ی اقلیدسی به‌کار می‌برند.

در این پایان‌نامه معیاری از اختلاف مرتبه‌ی اول مشاهدات که به‌صورت $w_{ij} = a_i - a_j$ به‌ازای $i,j = ۱, \dots, n$ است و در محاسبه‌ی واریانس به‌کار برده می‌شود، در فصل ۱ معرفی شده است. نتیجه‌ی اصلی، آماره‌ای برای اندازه‌گیری فاصله‌ی بین نمونه‌ای و روشی از تقسیم‌بندی پراکندگی کل است که در فصل ۲ معرفی می‌شود، ویژگی این آماره‌ها و معرفی آماره‌ی آزمون مؤلفه‌ی فاصله برای فرض کلی (۵) در فصل ۳ بیان می‌شود وتجزیه‌ی مؤلفه‌ی فاصله برای مدل‌های چندعاملی در فصل ۴ ذکر می‌شود. برای ارائه‌ی ویژگی سودمند این پایان‌نامه مثال‌ها و نتایج تجربی در فصل ۵ آورده شده‌اند.

---

[۲۷] Smouse

[۲۸] Quattro

[۲۹] McArdle

[۳۰] Zapala

[۳۱] Schork

فصل دوم

مؤلفه‌ی فاصله

۱-۱ مقدمه

از آن‌جایی‌که روش معمول محاسبه‌ی واریانس نمونه وابسته به اختلاف حول میانگین نمونه است، ممکن است فاقد صراحت باشد، به‌ویژه هنگامی‌که برنامه‌های رایانه‌ای برای محاسبه مورد استفاده قرار می‌گیرد. در این فصل روش‌های مختلف محاسبه‌ی واریانس ارائه می‌شود. روش پیشنهادی برای رفع مشکل محاسبه‌ی واریانس با استفاده از اختلاف حول میانگین، روش اختلاف مرتبه‌ی اول مشاهدات است، که در ابتدای فصل بیان می‌شود، در ادامه چندین مثال برای نشان دادن کاربرد این روش‌ها ارائه می‌شود. در پایان شیوه‌ی دیگر محاسبه‌ی واریانس با استفاده از فراوانی مشاهدات شرح داده خواهد شد.

۲-۱ محاسبه‌ی واریانس با استفاده از اختلاف مرتبه‌ی اول مشاهدات

فرض کنید $W = (w_{ij})$ یک ماتریس مثلثی با عناصر $w_{ij} = a_i - a_j \ (i, j = ۱,۲, \dots, n, i > j)$ به‌طوری‌که عناصر روی قطر اصلی صفر و عناصر بالای مثلث $w_{ij} (i > j)$ است. از روی این تعریف می‌توان به‌سادگی ملاحظه کرد که عناصر پایین قطر، از جمع عناصر $w_{j+۱\ j}, \dots, w_{i\ i-۱}$ در ماتریس W به‌دست می‌آید. برای مثال با یک نمونه به حجم $n = ۵$، اگر اختلاف مرتبه‌ی اول مشاهدات را به‌دست آوریم، داریم

$w_{۲۱} = d_{۱},\ w_{۳۲} = d_{۲},\ w_{۴۳} = d_{۳},\ w_{۵۴} = d_{۴}$

آن‌گاه

$$
\begin{aligned}
&w_{۳۱} = w_{۲۱} + w_{۳۲} = d_{۱} + d_{۲},\\
&w_{۴۱} = w_{۲۱} + w_{۳۲} + w_{۴۳} = d_{۱} + d_{۲} + d_{۳},\\
&w_{۴۲} = w_{۳۲} + w_{۴۳} = d_{۲} + d_{۳},\\
&w_{۵۱} = \textstyle\sum_{i=۱}^{n} d_i,\\
&w_{۵۲} = d_{۲} + d_{۳} + d_{۴},\\
&w_{۵۳} = d_{۳} + d_{۴}
\end{aligned}
$$

به‌طورکلی فرض کنید $w_{i+۱\,i} = d_i (i = ۱,۲,\dots,n)$ عناصر $l\ (l = ۱,۲,\dots,n)$ ام بالای قطر ماتریس W هستند، بنابراین

$$
\begin{aligned}
w_{i+l\,i} &= a_{i+l} - a_i\\
&= (a_{i+l} - a_{i+l-۱}) + (a_{i+l} - a_{i+l-۲})\\
&\quad + \cdots + (a_{i+۲} - a_{i+۱}) + (a_{i+۱} - a_i)\\
&= d_{i+l-۱} + d_{i+l-۲} + \cdots + d_{i+۱} + d_i\\
&= \textstyle\sum_{j=i}^{i+l-۱} d_j\,, \qquad\qquad i = ۱,۲,\dots,n-۱
\end{aligned}
$$

مثال ۱

برای محاسبه‌ی واریانس نمونه‌ی متشکل از $\left(۱۰۴,۹۴,۹۵,۱۰۱,۱۱۱\right)$، مشاهدات را به‌ترتیب $a_{۱} = ۹۴, a_{۲} = ۹۵, a_{۳} = ۱۰۱, a_{۴} = ۱۰۴, a_{۵} = ۱۱۱$ مرتب می‌کنیم. بنابراین مقادیر $w_{ij}\ (i,j = ۱,\dots,۵\ i > j)$ به‌صورت زیر به‌دست می‌آید:

$W_{۲۱} = ۱, W_{۳۱} = ۷, W_{۳۲} = ۶, W_{۴۱} = ۱۰, W_{۴۲} = ۹, W_{۴۳} = ۳,$

$W_{۵۱} = ۱۷, W_{۵۲} = ۱۶, W_{۵۳} = ۱۰, W_{۵۴} = ۷$

بنابراین

$$\textstyle\sum_{i=۲}^{۵}\sum_{j=۱}^{۴} w_{ij}^{۲} = ۱^{۲} + ۷^{۲} + ۶^{۲} + \cdots + ۷^{۲} = ۹۷۰,$$

در نتیجه با استفاده از رابطه‌ی(۳) داریم:

$$s_۵^۲ = \frac{۱}{۵ \times ۴} \times ۹۷۰ = \frac{۹۷۰}{۵ \times ۴} = ۴۸/۵.$$

قضیه ۱

فرض کنید d_i اختلاف مرتبه‌ی اول مشاهدات و $w_{ij} = \sum_{k=j}^{i-۱} d_k\ ;\ i > j$ باشد، آن‌گاه برای $n \geq ۲$ مشاهدات، واریانس با ۰

$$s_n^۲ = \frac{۱}{n(n-۱)} \boldsymbol{d}^T \boldsymbol{C} \boldsymbol{d}$$

نمایش داده می‌شود، که در آن $\boldsymbol{d} = (d_۱, d_۲, \dots, d_{n-۱})^T$ و $\boldsymbol{C} = (c_{ij})$ یک ماتریس متقارن $(n-۱) \times (n-۱)$ با $c_{ij} = c_{ji} = (n-i)j$ برای $i, j = ۱,۲, \dots, n, i \geq j$ است.

برهان:

از رابطه‌ی (۱) داریم $s_n^۲ = \frac{۱}{(n-۱)} TSS$ و از (۲) داریم:

$$\begin{aligned} nTSS &= \sum_{i=۲}^{n} \sum_{j=۱}^{i-۱} w_{ij}^۲ \\ &= \sum_{i=۲}^{n} \sum_{j=۱}^{i-۱} \sum_{k=j}^{i-۱} \sum_{l=j}^{i-۱} d_k d_l \\ &= \sum_{i=۲}^{n} \sum_{k=۱}^{i-۱} \sum_{j=۱}^{k} \sum_{l=j}^{i-۱} d_k d_l \\ &= \sum_{k=۱}^{n-۱} \sum_{i=k+۱}^{n} \sum_{l=۱}^{i-۱} \sum_{j=۱}^{min\,(k,l)} d_k d_l \\ &= \sum_{k=۱}^{n-۱} \sum_{l=۱}^{n-۱} \sum_{i=\max\,(k+۱,l+۱)}^{n} \sum_{j=۱}^{min\,(k,l)} d_k d_l \end{aligned}$$

اما

$$\sum_{i=\max\,(k+۱,l+۱)}^{n} \sum_{j=۱}^{min\,(k,l)} ۱ = \begin{cases} l(n-k) = c_{kl} & k \geq l \\ k(n-l) = c_{lk} & l \geq k \end{cases}$$

که در آن $c_{kl} = c_{lk}$. بنابراین

$$nTSS = \sum_{k=۱}^{n-۱} \sum_{l=۱}^{n-۱} d_k c_{kl} d_l = \boldsymbol{d}^T \boldsymbol{C} \boldsymbol{d}$$

واثبات کامل می گردد.

نتیجه ۱

برای یک نمونه به حجم $n \geq ۲$، رابطه‌ی بازگشتی زیر برقرار است:

$$\begin{aligned}&(n+۱)TSS(n+۱)-nTSS(n)\\&\quad=\left(\sum_{i=۱}^{n} d_i\right)^۲+\left(\sum_{i=۲}^{n} d_i\right)^۲+\cdots(d_{n-۲}+d_{n-۱}+d_n)^۲\\&\qquad+(d_{n-۱}+d_n)^۲+d_n^۲\end{aligned}$$

که در آن $TSS(m)$ مجموع کل مربعات m مشاهده منهای میانگین نمونه است.

۱-۳ مثالی برای محاسبه‌ی واریانس با استفاده از اختلاف مرتبه‌ی اول مشاهدات

مثال ۲

در محاسبه‌ی واریانس نمونه‌ای متشکل از $\left(۱۱۱, ۱۰۱, ۹۵, ۹۴, ۱۰۴\right)$ از $d_۱ = ۱, d_۲ = ۶$ $d_۳ = ۳, d_۴ = ۷$ استفاده می‌کنیم و به روش‌های زیرعمل می‌کنیم،

۱)محاسبه‌ی واریانس با استفاده از اختلاف مرتبه‌ی اول و مجموع‌شان:

طبق رابطه‌ی (۳) داریم:

$$\begin{aligned}s_۵^۲&=\frac{۱}{n(n-۱)}\sum_{i=۱}^{n}\sum_{j=۱}^{i-۱} w_{ij}^۲\\&=\frac{۱}{۲۰}(۱^۲+۶^۲+۳^۲+۷^۲)+\left(۷^۲+۹^۲+۱۰^۲\right)+\left(۱۰^۲+۱۶^۲\right)+۱۷^۲\\&=\frac{۹۷}{۲}=۴۸/۵\end{aligned}$$

۲) محاسبه‌ی واریانس با استفاده از اختلاف مرتبه‌ی اول و ماتریس ثابت $\boldsymbol{C}$:

از قضیه‌ی ۱ با $\boldsymbol{d} = (۱,۳,۶,۷)^T$ و

$$\boldsymbol{C} = \begin{bmatrix} ۱ & ۲ & ۳ & ۴ \\ ۳ & ۶ & ۴ & ۲ \\ ۲ & ۴ & ۶ & ۳ \\ ۱ & ۲ & ۳ & ۴ \end{bmatrix}$$

داریم:

$$s_۵^۲ = \frac{۱}{۲۰}\boldsymbol{d}^T\boldsymbol{C}\boldsymbol{d} = \frac{۹۷}{۲} = ۴۸/۵$$

۳) محاسبه‌ی واریانس از طریق رابطه‌ی بازگشتی:

از نتیجه‌ی ۱ داریم:

$$\begin{aligned} ۲TSS(۲) &= d_۱^۲ = ۱, \\ ۳TSS(۳) &= ۲TSS(۲) + (d_۱ + d_۲)^۲ + d_۲^۲ = ۱^۲ + ۷^۲ + ۶^۲ = ۸۶, \\ ۴TSS(۴) &= ۳TSS(۳) + (d_۱ + d_۲ + d_۳)^۲ + (d_۲ + d_۳)^۲ + d_۳^۲ \\ &= ۸۶ + ۱۰^۲ + ۹^۲ + ۳^۲ = ۲۷۶, \\ ۵TSS(۵) &= ۴TSS(۴) + (d_۱ + d_۲ + d_۳ + d_۴)^۲ + (d_۲ + d_۳ + d_۴)^۲ \\ &\quad + (d_۳ + d_۴)^۲ + d_۴^۲ = ۲۷۶ + ۱۷^۲ + ۱۶^۲ + ۱۰^۲ + ۷^۲. \\ &= ۹۷۰ \end{aligned}$$

بنابراین

$$s_۵^۲ = \frac{۱}{۲۰}۵TSS(۵) = \frac{۹۷}{۲} = ۴۸/۵$$

همان‌طور که ملاحظه می‌شود واریانس نمونه‌ای به‌دست آمده از هر سه روش ارائه شده یکسان است.

۱-٤ محاسبه‌ی واریانس مشاهدات با استفاده از فراوانی

واریانس یک نمونه طبقه‌بندی شده با k طبقه در یک توزیع فراوانی با مقادیرمیانی $y_۱, y_۲, \dots, y_k$ و فراوانی $f_۱, f_۲, \dots, f_k$ با