باسمه تعالی

تحلیل آماری

رشته:

مدیریت فناوری اطلاعات

دوره:

کارشناسی ارشد

تعداد واحد:

۲ واحد نظری

تهیه کننده:

هادی جباری نوقابی

عضو هیأت علمی‌دانشگاه فردوسی مشهد

پست الکترونیکی: Jabbarinh@um.ac.ir

درگاه: http://www.um.ac.ir/~jabbarinh

شهریورماه ۱۳۹۰

مقدمه:

روش‌های تحلیل و استنباط آماری از لحاظ فنی به دو دسته‌ی آمار پارامتری و آمار ناپارامتری تقسیم می‌شوند. روش‌های آمار پارامتری روش‌هایی هستند که مستلزم معلوم بودن توزیع جامعه هستند. بنابراین در صورتی که توزیع جامعه از الگوی خاصی مثل توزیع نرمال پیروی کند، می‌توان از این روش‌ها استفاده کرد. اما روش‌های آمار ناپارامتری آزاد توزیع بوده و به الگوی جامعه بستگی ندارند. در این درس ابتدا در مورد روش‌های استنباط پارامتری آماری به ویژه نحوه طراحی آزمایش‌ها و تحلیل نتایج آن‌ها مطالبی مطرح می‌شود. سپس مدل‌سازی و آنالیز رگرسیون توضیح داده می‌شود. در پایان نیز مبحث آزمون‌های ناپارامتری ارایه خواهد شد. در همه موضوعات این درس انتظار می‌رود دانشجو بتواند ضمن یادگیری مفاهیم آماری بتواند با نرم افزار آماری SPSS نیز تحلیل‌های آماری را انجام دهد. بنابراین هدف اصلی این درس آشنایی دانشجویان با مبانی طرح آزمایش‌ها، مدل‌سازی و رگرسیون و انجام استنباط‌های ناپارامتری آماری است.

مباحث این درس طبق سرفصل موجود در قالب شانزده جلسه تنظیم شده است که در هر جلسه دانشجویان با موضوعات آن جلسه آشنا می‌شوند.

فهرست مطالب

جلسه اول

۱. مبانی استنباط آماری، مفهوم و اهداف آن
۲. روش‌های استنباط آماری (برآورد نقطه‌ای، برآورد فاصله‌ای، آزمون فرضیه‌های آماری)
۳. استنباط نقطه‌ای درباره میانگین، نسبت، واریانس و ...

روش‌های آماری برای تجزیه و تحلیل داده‌ها به دو دسته‌ی کلی روش‌های آمار توصیفی و روش‌های آمار استنباطی دسته‌بندی می‌شوند. روش‌های آمار توصیفی شامل روش‌های جمع‌آوری داده‌ها، پردازش اولیه داده‌ها، دسته‌بندی و خلاصه کردن و همچنین رسم نمودارها و جداول مختلف آماری به همراه محاسبه شاخص‌های آماری تمایل مرکزی و پراکندگی است. در دروس قبلی با مفاهیم اولیه آمار و همچنین مسائل مرتبط با آمار توصیفی آشنا شده‌اید. در این درس به تشریح گزیده‌ای از روش‌های آمار استنباطی و کاربرد آن‌ها در مسایل مختلف علوم انسانی می‌پردازیم.

اصولاً استنباط آماری مبتنی بر تجزیه و تحلیل داده‌های جمع آوری شده از روش‌های نمونه‌گیری است. با استفاده از روش‌های مختلف استنباط آماری می‌توان روابط مختلف بین متغیرهای وابسته و مستقل را بررسی نموده و از روی این روابط و یا برازش مدل‌های مختلف به داده‌ها برای آینده نتیجه‌گیری نمود. همچنین با استفاده از این روش‌ها و با در نظر گرفتن سطح اطمینان معین، می‌توان نتایج بدست آمده از نمونه تصادفی را به جامعه تعمیم داد.

روش‌های آمار استنباطی عمدتاً شامل آزمون فرضیه‌های آماری و برازش مدل به داده‌ها است. گاهی اوقات نیز می‌توان روش برآورد فاصله‌ای را به عنوان یکی از روش‌های آمار استنباطی به شمار آورد.

وقتی براساس نمونه تصادفی تقریبی از میانگین، انحراف استاندارد یا واریانس و دیگر شاخص‌های آماری بدست می‌آوریم، در واقع برآورد نقطه‌ای از آن شاخص‌ها بر اساس نمونه تصادفی را محاسبه می‌کنیم.

واضح است که اگر نمونه تصادفی تغییر یابد، این برآوردها نیز عوض خواهند شد. لذا اگر بخواهیم با داشتن فقط یک نمونه تصادفی در مورد جامعه آماری و شاخص‌های آن استنباطی داشته باشیم، می‌توان از روش برآورد دیگری با نام برآورد فاصله‌ای استفاده نمود. در برآورد فاصله‌ای در واقع با درنظر گرفتن یک سطح اطمینان معین، فاصله‌ای را به عنوان حد بالا و پایین برای مقدار شاخص آماری در جامعه ارایه می‌کنیم.

مسأله آزمون فرضیه‌ها در اکثر مباحث آماری به عنوان یک مسأله اساسی مورد استفاده واقع می‌شود. در بیشتر مسایل، یک تناظر یک به یک بین آزمون فرضیه و برآورد فاصله‌ای پارامتر متناظر در جامعه وجود دارد.

انواع فرضیه‌ها

- فرضیه آماری H_0
- فرضیه مقابل H_1

معمولاً در آزمون فرضیه‌های آماری به بررسی ادعای محقق در مورد مقدار یا مقادیر پارامتر جامعه می‌پردازیم. لذا در آزمون فرضیه‌ها دو فرضیه به نام‌های فرضیه صفر یا اولیه که معمولاً با H_0 نشان داده می‌شود و فرضیه مقابل یا مخالف که با نماد H_1 نشان داده می‌شود، وجود دارد. فرضیه آماری هرگونه ادعایی در خصوص جامعه است که هیچ نوع رابطه یا تفاوتی را بیان نمی‌کند. فرضیه مقابل هر گونه فرضیه‌ای است که نفی H_0 باشد. بنابراین در اکثر موارد ادعای محقق تحت عنوان فرضیه مقابل درنظر گرفته می‌شود.

برای بررسی این فرضیه‌ها، با توجه به‌این که پارامتر مربوطه در جامعه نامعلوم است، باید نمونه‌ای تصادفی از جامعه برگزیده و با توجه به آماره‌های معین که بر اساس مشاهدات نمونه‌ای مقدار آن به دست می‌آید در

مورد رد کردن یا رد نکردن فرضیه آماری تصمیم‌گیری نمود. به طور کلی یک فرضیه گزاره‌ای در مورد پارامتر مجهول جامعه است.

مثال‌های زیر نمونه‌هایی از فرضیه‌های مقابل است:

۱- محققی ادعا می‌کند که متوسط ضریب هوشی دانشجویان علوم مهندسی بیشتر از ۱۱۰ است.

۲- پزشکی مدعی است که شیوع سرطان روده در مقایسه با دهه قبل افزایش یافته است.

۳- معلمی ادعا دارد که روش جدید تدریس ریاضی یادگیری بهتری را نسبت به روش‌های مشابه قبلی دارد.

به جهت درک بیشتر تناظر بین آزمون فرضیه‌ها و برآورد فاصله‌ای به مثال زیر توجه کنید:

فرض کنید می‌خواهیم برای پارامتر میانگین جامعه یک برآورد فاصله‌ای بدست آوریم. در این صورت حد وسط فاصله اطمینان، برآورد نقطه‌ای پارامتر میانگین جامعه است. حال اگر آزمون فرضیه به صورتی باشد که تحت فرضیه اولیه، پارمتر میانگین برابر مقداری معلوم باشد و فرضیه مقابل، نقیض آن باشد، با توجه به نمونه تصادفی از جامعه می‌توان به بررسی فرضیه‌ها پرداخت. اگر براساس نمونه، میانگین داده‌ها مقداری باشد که در حدود اطمینان برای پارمتر مجهول جامعه واقع گردد، فرضیه اولیه رد نمی‌شود. در غیر این صورت باید فرضیه اولیه را رد کنیم.

انواع خطاها

در آزمون فرضیه‌های آماری در دو حالت ممکن است که مرتکب خطا گردیم. اولین خطا که به خطای نوع اول مشهور است، وقتی رخ می‌دهد که فرضیه اولیه‌ای که واقعاً درست است را به اشتباه با توجه به نتایج نمونه رد کنیم. احتمال ارتکاب به خطای نوع اول را با نماد آلفا (α) نشان می‌دهیم. خطای بعدی که

به عنوان خطای نوع دوم شناخته می‌شود، وقتی رخ می‌دهد که فرضیه‌ی مقابل که واقعاً درست است را به اشتباه رد کنیم. احتمال ارتکاب به‌این خطا را با نماد بتا (β) نشان می‌دهیم. بنابراین داریم:

$\alpha =$ احتمال ارتکاب خطای نوع اول $= \mathrm{P}($ رد H_0 | H_0 درست باشد $)$

$\beta =$ احتمال ارتکاب خطای نوع دوم $= \mathrm{P}($ رد H_1 | H_1 درست باشد $)$

تذکر مهم

ارتکاب خطای نوع اول در مسایل آزمون فرضیه به مراتب بدتر از خطای نوع دوم است. به عنوان مثال فرض کنید، فردی متهم به یک نوع خلاف است. اگر در دادگاه، قاضی با توجه به تحقیقات به عمل آمده و مستندات پرونده، رای به مجرمیت این فرد بدهد در حالی که واقعا فرد گناهکار نیست، این مسأله اشتباه خیلی بدتری است نسبت به حالتی که فرد واقعاً گنهکار است، اما قاضی با توجه به شواهد موجود او را تبرئه نماید.

آماره آزمون

حال که مسأله آزمون فرضیه تشریح شد، می‌خواهیم بدانیم که چگونه از روی نمونه تصادفی در مورد فرضیه‌ها تصمیم‌گیری نمائیم. یکی از مشخصه‌هایی که به ما برای تصمیم‌گیری کمک می‌کند، آماره (آزمون) است. آماره تابعی از مشاهدات است که به پارامتر مجهول بستگی ندارد. بنابراین با داشتن یک نمونه تصادفی می‌توان مقدار آماره را تعیین نمود و از روی آن برای رد کردن یا عدم رد فرضیه اولیه در مقابل فرضیه مقابل تصمیم‌گیری کرد.

ناحیه بحرانی

ناحیه بحرانی آزمون مجموعه‌ای است که به ازای تمامی‌مقادیر آن، فرضیه اولیه رد می‌شود. بنابراین ناحیه بحرانی مقادیری از آماره آزمون است که به ازای آن فرضیه H_0 رد می‌شود.

سطح معنی‌داری آزمون

احتمال رد فرضیه اولیه وقتی در واقع آن فرضیه درست است، را سطح معنی‌داری آزمون نیز می‌نامند. معمولاً در تحقیقات تجربی یا شبه تجربی که احتمال ارتکاب خطای نوع اول به صورت اختیاری در دست محقق می‌باشد، مقدار سطح معنی‌داری آزمون برابر ۵ یا یک درصد در نظر گرفته می‌شود.

آزمون‌های یک طرفه و دو طرفه

با توجه به ناحیه رد و نوع فرضیه مقابل، آزمون فرضیه‌های آماری به دو دسته آزمون یک طرفه و دو طرفه تقسیم بندی می‌شود. آزمون‌های یک طرفه آزمون‌هایی هستند که ناحیه بحرانی آن‌ها فقط در یک طرف (راست یا چپ) توزیع احتمال آماره آزمون واقع می‌شود. آزمون‌هایی که ناحیه رد آن‌ها در دو طرف راست و چپ قرار می‌گیرد، آزمون‌های دو طرفه نامیده می‌شوند. آزمون‌های یک طرفه نیز به دو دسته آزمون‌های یک طرفه راست و یک طرفه چپ دسته بندی می‌شوند. اگر ناحیه رد یک آزمون یک طرفه در سمت راست باشد، به آن آزمون یک طرفه راست یا بالا دنباله‌ای گفته می‌شود و در غیر این صورت آن را یک طرفه چپ یا پائین دنباله‌ای می‌نامند.

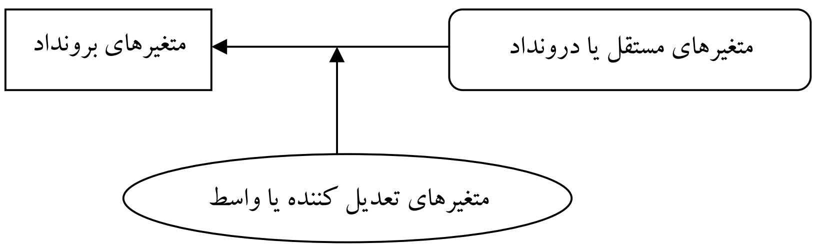
اگر محقق در ادعای خود تأکید بر نوعی برتری روش جدید نسبت به وضع موجود داشته باشد، همانند این که روش جدید آموزشی نسبت به روش فعلی کاراتر است، آن‌گاه آزمون فرضیه مربوطه یک طرفه خواهد بود. به عنوان مثال فرض کنید می‌خواهیم $H_0: \mu \geq 50$ را در مقابل $H_1: \mu < 50$ آزمون کنیم. چون تنها به ازای مقادیر کوچک میانگین جامعه، فرضیه اولیه رد می‌شود و ناحیه رد آزمون در طرف چپ (از لحاظ محور اعداد) واقع می‌شود، این آزمون، یک آزمون فرضیه یک طرفه چپ است. اما اگر $H_0: \mu \leq 50$ باشد، باید فرضیه مقابل را به صورت $H_1: \mu > 50$ در نظر گیریم. در این صورت چون ناحیه بحرانی در سمت راست μ قرار می‌گیرد، آزمون فوق از نوع یک طرفه راست است.

جلسه دوم

۱. برازش مدل به داده‌ها
۲. مفاهیم مقدماتی طرح آزمایش‌ها
۳. مفهوم مقایسه میانگین دو یا چند جامعه مستقل
۴. تحلیل واریانس یک طرفه و طرح کاملاً تصادفی

مقدمه

در بسیاری از مسایل کاربردی مدل‌سازی نقش مهمی‌در جهت توصیف ارتباط بین متغیرها و پیش بینی برای آینده دارد. منظور از مدل سازی، مطالعه‌ی روابط و اثرات متغیرها روی یکدیگر است. در واقع در مدل‌سازی به دنبال برازش یک الگوی فرضی به داده‌ها هستیم. این الگوی فرضی باید به پیشنهاد دانشمندان مختلف و مبتنی بر اصول و منطق علمی‌بین پدیده‌ها باشد. شکل زیر نمونه‌ای از یک مدل مفهومی‌برای توصیف ارتباط بین متغیرها است.



شکل ۱، نمونه‌ای از یک مدل مفهومی‌ارتباط بین متغیرها

در بسیاری از مطالعات نیمه‌تجربی مهم‌ترین راه جمع‌آوری داده‌ها، انجام آزمایش‌های طرح‌ریزی شده است. برای تجزیه و تحلیل اطلاعات حاصل از این آزمایش‌ها نیز از نوعی مدل‌سازی استفاده می‌شود. منظور از طرح آزمایش، تدوین و طراحی یک نقشه برای جمع‌آوری داده‌ها و مطالعات اثرات متغیرها روی یکدیگر است. طرح آزمایش را می‌توان با نقشه‌های ساختمانی که یک مهندس ساختمان برای احداث یک واحد مسکونی ارایه می‌دهد، مقایسه نمود. با وجودی که مهندس می‌تواند با بکارگیری خلاقیت خود نقشه‌های متنوعی را ارایه دهد، اما موظف است که نیازهای اساسی ساکنین آینده ساختمان یا کارفرما را برآورده سازد. برای این منظور، چندین طرح مختلف ارایه کرده و از بین آن‌ها با توجه به تمام جوانب امر طرح موردنظر را انتخاب می‌کند. در طراحی یک آزمایش، طراح در نقش مهندس و آزمایش‌گر در نقش صاحب ساختمان یا کارفرما است و اوست که تصمیم‌گیری نهایی را درباره‌ی آزمایش اتخاذ می‌کند. پس

بهترین فرد برای یک آزمایش کسی است که

الف) بیش‌ترین شناخت را با ماهیت مواد آزمایشی داشته باشد.

ب) بالاترین توانایی را برای ارزشیابی محاسن و معایب روش‌های مختلف دارا باشد.

ج) بیشترین آشنایی را با روش‌های مختلف ممکن برای طرح‌ریزی آزمایش داشته باشد.

بدون وجود یک طرح آزمایش مناسب نمی‌توان فرضیه‌های بالقوه سودمند را با درجه قابل قبولی از دقت آزمایش کرد. قبل از رد کردن فرضیه‌ای مربوط به یک زمینه‌ی تحقیقاتی، باید ساختار آزمایش را مورد بررسی قرار داد و اطمینان حاصل کرد که آیا آزمایش یک آزمون واقعی از فرضیه را فراهم ساخته یا خیر؟ به معنای واقعی، آزمایش یک آزمون است. آزمایش طرح‌ریزی شده، یک آزمون یا دنباله‌ای از آزمون‌ها است که در آن‌ها تغییراتی موردنظر در متغیرهای ورودی فرآیند یا سیستم اعمال می‌شود به قسمی که بتوانیم علل تغییرات در پاسخ خروجی را مشاهده و مشخص کنیم. اهداف آزمون می‌تواند شامل موارد ذیل باشد:

۱. تعیین متغیرهایی که بیش‌ترین تأثیر را در پاسخ و نتیجه‌ی آزمایش دارد.
۲. تعیین موقعیت متغیرهای مؤثر طوری که تقریباً پاسخ همیشه نزدیک مقدار مطلوب اسمی‌باقی بماند.
۳. تعیین موقعیت متغیرهای مؤثر طوری که تغییر پذیری متغیر پاسخ کوچک باشد.
۴. تعیین موقعیت متغیرهای مؤثر طوری که اثرهای کنترل ناپذیر کمینه شوند.

انواع آزمون‌های آماری

- **آزمون‌های پارامتری**
  - آزمون استیودنت
    - مقایسه‌ی میانگین جامعه با عدد ثابت ($H_0 : \mu = \mu_0$)
    - مقایسه‌ی میانگین دو جامعه وابسته ($H_0 : \mu_1 = \mu_2$)
    - مقایسه‌ی میانگین دو جامعه مستقل ($H_0 : \mu_1 = \mu_2$)
  - تحلیل واریانس بین گروهی یک یا چند طرفه
    - مقایسه‌ی میانگین سه یا چند جامعه مستقل (آنالیز واریانس یک طرفه‌ی پارامتری $H_0 : \mu_1 = \mu_2 = ... = \mu_k$)
    - آنالیز واریانس دو یا چند طرفه‌ی پارامتری
  - تحلیل واریانس درون گروهی یک یا چند طرفه
    - آنالیز طرح اندازه‌های تکراری

- آزمون‌های همبستگی
  - آزمون معنی‌داری ضریب همبستگی خطی پیرسن
  - تحلیل رگرسیون ساده
  - تحلیل رگرسیون چندگانه

– **آزمون‌های ناپارامتری**

- آزمون مقایسه‌ی میانه‌ی جامعه با عدد ثابت
  - آزمون علامت تک نمونه‌ای
  - آزمون رتبه‌ی علامت‌دار ویلکاکسون تک نمونه‌ای
- آزمون مقایسه‌ی میانه‌ی دو جامعه وابسته
  - آزمون علامت دو نمونه‌ای
  - آزمون رتبه‌ی علامت‌دار ویلکاکسون دو نمونه‌ای
  - آزمون مک نمار
- آزمون مقایسه‌ی میانه‌ی دو جامعه مستقل
  - آزمون من–ویتنی–ویلکاکسون
  - آزمون میانه
- آزمون مقایسه‌ی میانه‌ی سه یا چند جامعه مستقل
  - آزمون کروسکال والیس
  - آزمون میانه
- آزمون مقایسه‌ی میانه‌ی سه یا چند جامعه وابسته
  - آزمون فریدمن
  - آزمون ککران
- آزمون‌های نیکویی برازش
  - آزمون نیکویی برازش کلموگروف–اسمیرنوف
  - آزمون نیکویی برازش پیرسن
- آزمون دورها (گشت‌ها)
- آزمون معنی‌داری ضریب همبستگی اسپیرمن

در دروس قبلی فرا گرفتید که برای میانگین مجهول یک جامعه می‌توان آزمون برابری میانگین با یک عدد ثابت را انجام داد. همچنین اگر دو جامعه مستقل یا وابسته را داشته باشیم، می‌توانستیم آزمون برابری

میانگین این دو جامعه را در برابر فرضیه‌ی مقابل یک طرفه یا دو طرفه انجام دهیم. به همین ترتیب اگر بخواهیم در مورد پارامتر پراکندگی یعنی واریانس یک جامعه، استنباطی داشته باشیم می‌توانیم با بنا نهادن یک آزمون فرضیه، برابری واریانس جامعه با عدد معلوم را بررسی کنیم. اگر دو جامعه مستقل نیز وجود داشته باشد، قبلاً مشاهده کردیم که آزمون برابری واریانس دو جامعه یا معادل آن آزمون برابری نسبت واریانس دو جامعه با عدد یک، نیز با در نظر گرفتن آماره مرتبط قابل انجام است. به علاوه اگر بخواهیم نسبت یک صفت مجهول را در جامعه‌ای بررسی نماییم، به راحتی می‌توان آزمون برابری نسبت با یک عدد معلوم را انجام داده یا این که نسبت یک صفت را در دو جامعه مستقل بررسی کنیم.
اما اگر بخواهیم پارامتر میانگین بیش از دو جامعه‌ی مستقل را با یکدیگر مقایسه کنیم یا این که این مقایسه در مواجهه با پارامترهای مزاحم دیگر بررسی شود، باز هم با استفاده از تکنیک‌های مدل‌سازی و آزمون آن می‌توان اثرات متغیرها روی یکدیگر را مطالعه نمود. در واقع در این‌گونه موارد باید یک مدل آماری به داده‌ها برزاش داده و بر اساس مدل بدست آمده، به آزمون فرضیه‌های مرتبط پرداخت.

آنالیز واریانس یک طرفه

- ***آماره آزمون***
- ***توزیع فیشر***
- ***برابری یا نابرابری حجم نمونه***
- ***فرضیه‌ها***
- ***نحوه‌ی انجام آزمون***
- ***درجه‌ی آزادی صورت و مخرج***
- ***نحوه‌ی تصمیم‌گیری***

در این‌جا می‌خواهیم آزمون فرضیه‌ی تساوی میانگین‌های بیش از دو جامعه را در مقابل فرضیه‌ی جانشین «عدم تساوی میانگین‌ها» انجام دهیم. به عنوان مثال، مایلیم براساس نمونه تصمیم‌گیری نماییم که آیا واقعاً تفاوتی بین سه روش تدریس یک زبان خارجی وجود دارد یا فرض کنید می‌خواهیم متوسط محصول شش نوع گندم را در هر هکتار با هم مقایسه کنیم یا این که می‌خواهیم میزان محصولات سه ماشین صنعتی که یک نوع کالای معین را تولید می‌کنند، مقایسه نماییم. در همه این مسایل یک عامل چند سطحی داریم و می‌خواهیم اثر آن عامل را روی متغیر پاسخ مطالعه کنیم. معمولاً در طرح آزمایش‌ها به سطوح عامل، تیمار هم گفته می‌شود. در این گونه طرح‌ها، تیمارها را به طور کاملاً تصادفی به مجموعه واحدهای آزمایش نسبت می‌دهیم و به این دلیل به آن‌ها طرح تک عاملی کاملاً تصادفی شده می‌گوییم. این طرح وقتی

کاراست که بین واحدهای آزمایش هیچ‌گونه منبع تغییر شناخته شده‌ی دیگری به جز تیمارها وجود نداشته باشد. برای بررسی مسایلی نظیر موارد فوق باید از تکنیکی به نام آنالیز واریانس یک طرفه استفاده شود. در این آزمون‌ها، فرض بر این است که جامعه‌ها دارای توزیع نرمال بوده و واریانس‌های مساوی داشته باشند. همچنین نمونه‌های تصادفی از جامعه‌ها باید مستقل باشند. در این روش، ابتدا واریانس کل را به دو مؤلفه‌ی پراکندگی بین گروه‌ها و داخل گروه‌ها تجزیه می‌کنیم. بنابراین به این روش آنالیز (تحلیل) واریانس گفته می‌شود. روش آنالیز واریانس را به اختصار با ANOVA نمایش می‌دهند که مخفف عبارت Analyze Of Variance است.

آماره آزمون

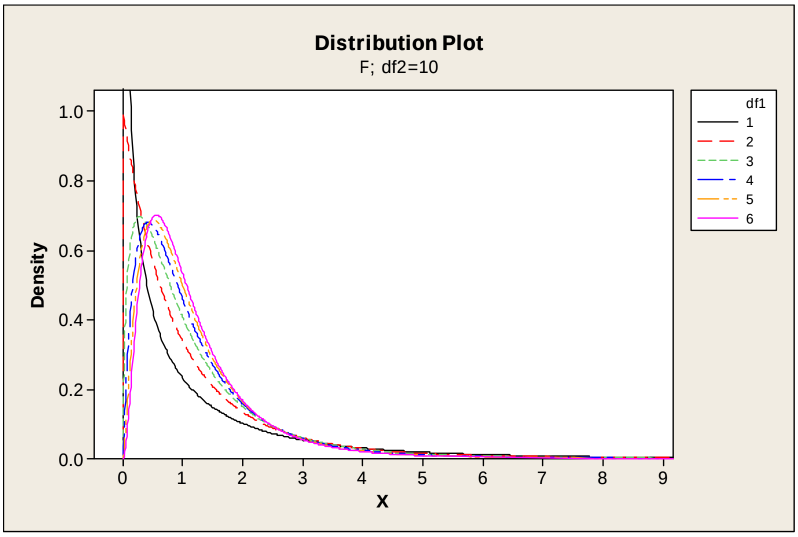
مطابق روش معمول در آزمون فرضیه‌ها، بایستی یک معیار تحت عنوان آماره آزمون معرفی گردد. به طوری که بعداً مشاهده می‌شود، آماره آزمون فوق دارای توزیع فیشر (F) است.

توزیع فیشر

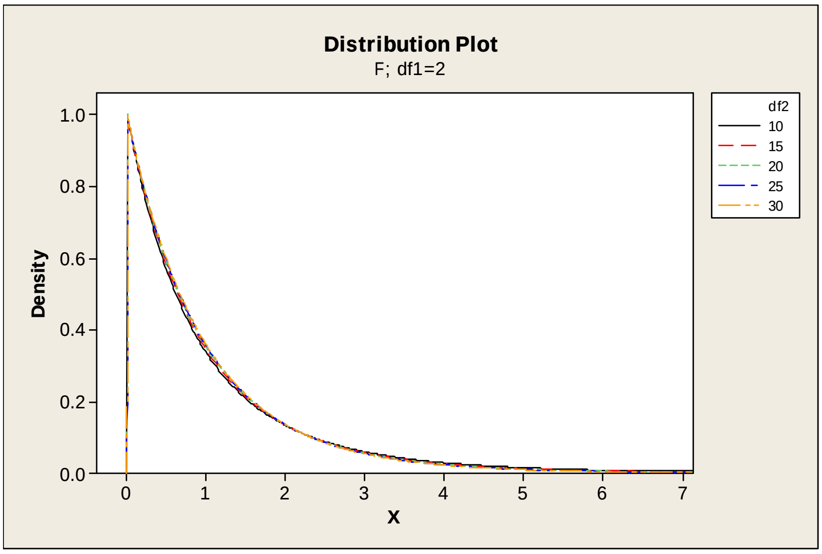
توزیع فیشر دارای ویژگی‌های زیر است.

۱. توزیعی است نامتقارن و دارای کجی یا چولگی راست
۲. مقادیر متغیر تصادفی توزیع فیشر نامنفی هستند.
۳. شکل توزیع فیشر با دو عدد مثبت که تحت عنوان درجه آزادی صورت و مخرج نام گذاری می‌شوند، مشخص می‌گردد.

نمونه‌هایی از توزیع فیشر با درجات مختلف آزادی صورت



نمونه‌هایی از توزیع فیشر با درجات مختلف آزادی صورت



تذکر

اگر یک طریق برای طبقه‌بندی جامعه‌ای در نظر گرفته شود، روشی که برای انجام آزمون فرضیه‌ی صفر «تساوی میانگین‌ها» به کار برده می‌شود را روش آنالیز واریانس یک طرفه و اگر موقعیت‌هایی وجود داشته باشد که دو راه برای طبقه بندی جامعه در نظر گرفته شود، آنالیز واریانس را دو طرفه یا دو عاملی نامند. در بررسی این موارد ممکن است نمونه‌هایی که از هر جامعه انتخاب می‌شوند دارای حجم‌های مساوی یا نامساوی باشند که آن‌ها را به تفکیک بیان خواهیم نمود.

آنالیز واریانس یک طرفه وقتی حجم نمونه‌ها مساوی باشد

در این جا آزمون فرضیه‌هایی که شامل بیش از دو میانگین مساوی هستند در مقابل فرضیه‌ی مخالف را در نظر می‌گیریم. یعنی برای $k \geq 1$ داریم:

$$\begin{cases} \mathrm{H}_0 : \mu_1 = \mu_2 = \ldots = \mu_k \\ \mathrm{H}_1 : \mu_1 \neq \mu_2 \neq \ldots \neq \mu_k \end{cases}$$

مسایلی که در این قسمت مطرح می‌شود، مواردی را شامل می‌شوند که حجم نمونه‌های منتخب از هر گروه، مساوی و تنها یک راه برای طبقه بندی جامعه وجود دارد.

روش آزمون

چنان که اشاره گردید، روش آنالیز واریانس مستلزم این است که دو برآورد متفاوت از واریانس مشترک جامعه‌ها (σ^2) بدست آورده و آن دو برآورد را به وسیله توزیع فیشر با هم مقایسه کنیم. آماره‌ی آزمون نسبت واریانس بین گروه‌ها به واریانس داخل گروه‌ها در نظر گرفته می‌شود.

واریانس بین گروه‌ها برآوردی از واریانس مشترک است که با توجه به میانگین نمونه‌ها بدست می‌آید. هدف این است که تغییراتی که بر اثر اختلاف بین میانگین گروه‌ها به وجود می‌آید و به اختلاف تیمارها منجر می‌شود را اندازه بگیریم. اگر S_x^2 را واریانس میانگین‌های نمونه‌ها در نظر بگیریم، خواهیم دید که واریانس بین گروه‌ها (نمونه‌ها) برابر nS_x^2 خواهد بود. اگر قدری تأمل کنیم، به این نتیجه دست خواهیم یافت که nS_x^2 می‌تواند برآوردی از σ^2 باشد، زیرا می‌دانیم $\sigma_x^2 = \frac{\sigma^2}{n}$ و از این جا نتیجه می‌شود که $\sigma^2 = n\sigma_x^2$ که بیان‌گر این است که می‌توان σ^2 را با nS_x^2 برآورد کرد.

واریانس داخل گروه‌ها که آن را تغییرات به واسطه‌ی خطا نیز می‌نامند، برآوردی از σ^2 براساس واریانس‌های نمونه‌ها است. اگر S_p^2 را واریانس مشترک یا ادغام شده بنامیم که برابر میانگین واریانس نمونه‌ها خواهد بود، آن‌گاه می‌توان گفت واریانس داخل نمونه‌ها برابر S_p^2 است. با توجه به آماره آزمون فیشر که قبلاً تعریف شد، می‌بینیم که هر چه میانگین‌های نمونه‌ها از هم فاصله بیشتری داشته باشند، مقدار صورت کسر آماره آزمون نیز بزرگ‌تر شده و در نتیجه مقدار F بزرگ‌تر خواهد شد.

درجات آزادی صورت و مخرج

آزمون F یک آزمون یک طرفه راست است، زیرا مقادیر خیلی بزرگ آماره نتیجه عدم تساوی میانگین‌ها می‌باشد. مقدار بحرانی F را از جدول توزیع فیشر می‌توان به دست آورد که در آن α سطح معنی‌داری آزمون بوده و درجات آزادی صورت و مخرج به صورت زیر می‌باشند:

(k تعداد گروه‌ها یا نمونه‌ها) $df_1=$درجه آزادی صورت$=k-1$

(n حجم هر یک از نمونه‌ها) $df_2=$درجه آزادی مخرج$=k(n-1)$

نحوه‌ی تصمیم‌گیری

اگر مقدار آماره آزمون از مقدار بحرانی جدول توزیع فیشر با درجه آزادی‌های $k-1$ و $k(n-1)$ بیشتر باشد، فرضیه‌ی صفر مبنی بر تساوی میانگین گروه‌ها، رد می‌شود. در غیر این صورت فرضیه‌ی اولیه رد نمی‌شود.

مثال ۱. می‌خواهیم میزان محصول سه ماشین صنعتی را با هم مقایسه کنیم. برای این منظور میزان محصول هر یک از ماشین‌ها را در پنج ساعت مختلف ثبت نموده و نتایج را در جدول زیر آورده‌ایم. در سطح معنی‌داری پنج درصد این ادعا را که متوسط تولید سه ماشین یکسان می‌باشد، آزمون کنید.

شماره ماشین	نمونه‌ها از ماشین i ام					$\overline{X}_i$
$i=$۱	۴۷	۵۳	۴۹	۵۰	۴۶	۴۹
$i=$۲	۵۵	۵۴	۵۸	۶۱	۵۲	۵۶
$i=$۳	۵۴	۵۰	۵۱	۵۱	۴۹	۵۱

در این جا برای بررسی ادعا باید آزمون فرضیه $H_0: \mu_1=\mu_2=\mu_3$ را در مقابل فرضیه H_1 «عدم تساوی میانگین‌ها» انجام دهیم. برای این منظور باید واریانس بین گروه‌ها و واریانس داخل گروه‌ها را پیدا کنیم. لذا داریم:

$$S_{\bar{x}}^2=\tfrac{1}{k-1}\sum_{i=1}^{k}\left(\bar{x}_i-\bar{x}\right)^2$$

که در آن $\bar{x}$ میانگین کل داده‌ها است. بنابراین

$$\bar{x}=\tfrac{47+53+...+55+...+54+...+49}{15}=52$$

و

$$S_{\bar{x}}^2=\tfrac{1}{3-1}\left[(49-52)^2+(56-52)^2+(51-52)^2\right]=13$$

لذا واریانس بین نمونه‌ها $nS_{\bar{x}}^2=5\times13=65$ می‌باشد. حال واریانس داخل نمونه‌ها را با توجه به S_p^2 و به صورت زیر می‌یابیم. برای این منظور ابتدا S_1^2، S_2^2 و S_3^2 را محاسبه می کنیم.

$$S_1^2=\tfrac{1}{n-1}\sum_{j=1}^{n}\left(x_{1j}-\bar{x}_1\right)^2=\tfrac{1}{5-1}\left[(47-49)^2+(53-49)^2+...+(46-49)^2\right]=7.5$$

$$S_2^2=\tfrac{1}{n-1}\sum_{j=1}^{n}\left(x_{2j}-\bar{x}_2\right)^2=\tfrac{1}{5-1}\left[(55-56)^2+(54-56)^2+...+(52-56)^2\right]=12.5$$

$$S_3^2=\tfrac{1}{n-1}\sum_{j=1}^{n}\left(x_{3j}-\bar{x}_3\right)^2=\tfrac{1}{5-1}\left[(54-51)^2+(50-51)^2+...+(49-51)^2\right]=3.5$$

لذا خواهیم داشت $S_p^2=\tfrac{S_1^2+S_2^2+S_3^2}{3}=\tfrac{7.5+12.5+3.5}{3}=7.83$ و واریانس داخل گروه‌ها برابر ۷/۸۳ می‌شود. حال مقدار آماره آزمون را می‌یابیم.

$$F=\frac{nS_{\bar{x}}^2}{S_p^2}=\tfrac{65}{7.83}=8.3$$

مقدار بحرانی F با درجات آزادی ۲ و ۱۲ به ترتیب برای صورت و مخرج با استفاده از جدول توزیع فیشر با سطح معنی‌داری ۵ درصد برابر ۳/۸۹ است. چون مقدار آماره‌ی آزمون از مقدار بحرانی بیشتر است، فرضیه‌ی صفر یعنی برابری میانگین‌های محصول سه ماشین در سطح خطای ۵ درصد رد می‌شود. ملاحظه می‌شود که تفاوت بین میانگین‌های نمونه‌ها نسبت به عدد معرف نوسانات، بسیار بزرگ است.

مثال ۲. فرض کنید می‌خواهیم سه نمونه را فقط از یک ماشین بدست آوریم (جدول زیر). همان طور که انتظار داریم اگر چه $\overline{X}_i$ها در این مورد یکی هستند ولی خطای نمونه‌گیری سبب تفاوت‌های کمی می‌شود. حال سؤال زیر را مطرح می‌کنیم:

آیا تفاوت‌های موجود در $\overline{X}_i$ های داده‌های مربوط به محصولات سه ماشین در همان حدود تفاوت‌های داده‌های انتخاب شده از محصولات یک ماشین است؟

شماره نمونه	مقادیر نمونه‌ها					$\overline{X}_i$
$i=۱$	۴۹	۵۵	۵۱	۵۲	۴۸	۵۱
$i=۲$	۵۲	۵۱	۵۵	۵۸	۴۹	۵۳
$i=۳$	۵۱	۵۱	۵۲	۵۲	۵۰	۵۱/۲

بر اساس داده‌های جدول فوق، مقدار $S_{\bar{x}}^2$ برابر ۱/۲۱۳ می‌باشد. لذا مقدار آماره آزمون عبارت است از:

$$F=\frac{6.067}{6.9}=0.879$$

چون مقدار آماره آزمون از مقدار بحرانی جدول توزیع فیشر یعنی ۳/۸۹ کمتر است، فرضیه‌ی اولیه را نمی‌توان رد نمود. در این حالت تفاوت‌های بین میانگین‌های نمونه‌ها را می‌توان به طور معقول با توسل به نوسانات شانسی تبیین کرد. این امر شگفت آور نیست زیرا ما سه نمونه موجود را از یک ماشین به دست آورده‌ایم.

مثال ۳. حال نمونه‌های جدول زیر را در نظر می‌گیریم که پراکندگی زیادتری دارند.

شماره ماشین	نمونه از ماشین iام					$\overline{X}_i$
$i=۱$	۵۷	۴۲	۵۳	۳۸	۵۵	۴۹
$i=۲$	۴۶	۵۹	۶۴	۶۱	۵۰	۵۴
$i=۳$	۵۷	۵۹	۴۸	۴۶	۴۵	۵۱

داده‌های جدید از ماشین‌های پرخطاتری هستند که نوسانات شانسی بزرگ‌تری در داخل هر ماشین تولید می‌کنند. مقدار آماره‌ی آزمون برای این داده‌ها به صورت زیر است:

$$F=\frac{5(13)}{57.5}=1.13$$

چون مقدار آماره‌ی آزمون یعنی ۱/۱۳ از مقدار صدک توزیع فیشر با ۲ و ۱۲ درجه آزادی و سطح معنی‌داری ۵ درصد کمتر است، فرضیه‌ی H_0 را نمی‌توان رد کرد. در این حالت اثر تفاوت بین میانگین‌های نمونه‌ها در نوسانات موجود در مخرج کسر محو می‌شود.

جدول آنالیز واریانس

در این قسمت روش متداول و مناسبی را جهت انجام آنالیز واریانس ارایه می‌کنیم. دوباره یادآوری می‌شود که فرض بر این است که تمام نمونه‌ها از جامعه‌های نرمال با واریانس یکسان σ^2 گرفته شده‌اند. ولی در عمل امکان دارد میانگین‌های این جامعه‌ها متفاوت باشند که در واقع این تفاوت محتمل بین میانگین‌ها است که مورد آزمون واقع می‌گردد. جدول زیر جدول آنالیز واریانس نام دارد که شامل خلاصه‌ی تمامی شاخص‌ها و آماره‌های بدست آمده از نمونه‌ها است و ستون آخر آن مقدار آماره‌ی F است.

منبع تغییرات	مقدار	درجه آزادی	میانگین مجموع توان‌های دوم یا واریانس	آماره‌ی F
بین گروه‌ها، تبیین شده توسط تفاوت‌های بین $\overline{X}_i$ ها	$SS_B = n\sum_{i=1}^{k}\left(\overline{X}_i - \overline{X}\right)^2$	$k-1$	$nS_{\bar{x}}^2$	$F = \frac{MS_B}{MS_w}$
داخل گروه‌ها، تغییرات باقیمانده حاصل از نوسانات شانسی "تبیین نشده"	$SS_w = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_i\right)^2$	$k(n-1)$	S_p^2	--
جمع	$SS_T = SS_B + SS_w$	$nk-1$	--	--

به طور کلی در جدول آنالیز واریانس می‌توان گفت که تغییرات کل برابر مجموع تغییرات بین گروه‌ها و داخل گروه‌ها می‌باشد.

مثال ۴. می‌خواهیم برای استفاده در رادیوی چراغ قوه خانگی، باتری مناسبی که قابل شارژ باشد، خریداری نماییم. سه نوع باتری در بازار موجود است. برای این که طول عمر باتری‌ها را مقایسه نماییم از هر نوع باتری پنج عدد خریداری نموده و زمان کارکرد آن‌ها را به نحو جدول زیر ثبت می‌کنیم. آیا سه نوع باتری دارای زمان کارکرد یکسانی هستند (آزمون را در سطح معنی‌داری پنج درصد انجام دهید)؟

نوع یک	۲۶	۲۸/۵	۲۷/۳	۲۵/۹	۲۸/۲
نوع دو	۲۹	۲۸/۸	۲۷/۶	۲۸/۱	۲۷
نوع سه	۳۰	۲۶/۳	۲۹/۲	۲۷/۱	۲۹/۸

باید فرضیه‌ی $H_0: \mu_1 = \mu_2 = \mu_3$ را در مقابل فرضیه‌ی عدم برابری طول عمر باتری‌ها بیازماییم. برای تشکیل جدول آنالیز واریانس، ابتدا باید تغییرات تبیین شده و نشده را پیدا کنیم. داریم:

$$\overline{X}_1 = 27.18\,,\ \overline{X}_2 = 28.10\,,\ \overline{X}_3 = 28.48\,,\ \overline{X} = \tfrac{27.18+28.10+28.48}{3} = 27.92$$

لذا

$$SS_B = 5\left[(27.18-27.92)^2 + (28.10-27.92)^2 + (28.48-27.92)^2\right] = 4.5$$

$$SS_T = \sum_{i=1}^{3}\sum_{j=1}^{5}\left(X_{ij}-\overline{X}\right)^2 = \left[(26-27.92)^2 + \ldots + (28.2-27.92)^2 + (29-27.92)^2 + \ldots + (27-27.92)^2\right.$$
$$\left. + (30-27.92)^2 + \ldots + (29.8-27.92)^2\right] = 23.3$$

بنابراین $SS_w = SS_T - SS_B = 23.3 - 4.5 = 18.8$ و جدول آنالیز واریانس به صورت زیر است.

منبع تغییرات	مقدار	درجه آزادی	میانگین مربعات یا واریانس	آماره F
تبیین شده	۴/۵	۲	۲/۲۵	۱/۳۶
تبیین نشده	۱۸/۸	۱۲	۱/۶۵	--
جمع	۲۳/۳	۱۴	--	--

مقدار بحرانی با درجه آزادی دو و دوازده و سطح معنی‌داری ۵ درصد برابر ۳/۸۹ است. لذا با مقایسه این مقدار بحرانی با آماره آزمون، نتیجه می‌گیریم که در سطح خطای ۵ درصد، فرضیه‌ی اولیه یعنی برابری طول عمر باتری‌های مختلف رد نمی‌شود.

آنالیز واریانس یک طرفه وقتی حجم نمونه‌ها مساوی نباشد

روش انجام آنالیز واریانس یک طرفه در حالتی که حجم نمونه‌ها یکسان نیست تقریباً همانند حالتی است که حجم نمونه‌ها برابر هستند. تنها تفاوت در محاسبه‌ی واریانس بین گروه‌ها و داخل آن‌ها می‌باشد. فرض می‌کنیم حجم نمونه‌های انتخابی از هر کدام از گروه‌های ۱، ۲،...و k ام به ترتیب n_1، n_2،... و n_k باشد که در آن $N = n_1 + n_2 + \ldots + n_k$. برای آزمون فرضیه‌ی برابری میانگین‌های گروه‌ها، همانند قبل باید پراکندگی بین گروه‌ها را نسبت به پراکندگی داخل گروه‌ها بسنجیم. برای این منظور واریانس بین گروه‌ها و داخل گروه را به صورت زیر محاسبه می‌کنیم:

واریانس بین گروه‌ها= $MS_B = \frac{1}{k-1}\sum_{i=1}^{k} n_i\left(\overline{X}_i - \overline{X}\right)^2$ یا $SS_B = \sum_{i=1}^{k} n_i\left(\overline{X}_i - \overline{X}\right)^2$

$$SS_w = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_i\right)^2 \quad \text{یا} \quad MS_w = \frac{\sum_{i=1}^{k}(n_i - 1)S_i^2}{\sum_{i=1}^{k}(n_i - 1)} = \text{واریانس داخل گروه‌ها}$$

در این جا چون حجم نمونه‌های انتخابی از هر گروه برابر نیست، درجه‌ی آزادی بین گروه‌ها برابر $k-1$ و درجه‌ی آزادی داخل گروه‌ها برابر $N-k$ است که در آن k تعداد گروه‌ها و N تعداد کل نمونه‌های انتخابی از همه گروه‌ها می‌باشد.

تذکر

اگر در فرمول‌های فوق قرار دهیم $n_1 = n_2 = \dots = n_k = n$ آن‌گاه فرمول‌های قبلی برای محاسبه‌ی واریانس داخل و بین گروه‌ها بدست خواهد آمد.

برای آنالیز واریانس یک طرفه وقتی حجم نمونه‌ها یکسان نیست، باید مقدار آماره‌ی آزمون را با توجه به فرمول‌های مربوطه محاسبه کنیم و سپس مقدار آماره‌ی آزمون را با مقدار بحرانی که از جدول توزیع فیشر با سطح معنی‌داری α و درجه‌های آزادی $k-1$ و $N-k$ به ترتیب برای صورت و مخرج مقایسه نماییم. اگر مقدار آماره‌ی آزمون بزرگ‌تر از مقدار بحرانی باشد، همانند قبل فرضیه‌ی تساوی میانگین‌های گروه‌ها را رد می‌کنیم.

مثال ۵. سه نوع گیربکس مختلف روی اتومبیل‌های مختلفی که همگی موتورهای شش سیلندر با حجم موتور ۳۰۰۰ میلی لیتر دارند، کار گذاشته شده است. مصرف بنزین به مایل بر گالن، نتایج زیر را دربر داشته است. در سطح معنی‌داری پنج درصد، این ادعا را آزمون کنید که متوسط میزان مصرف بنزین خودروهای با گیربکس‌های نوع مختلف، متفاوت نمی‌باشد.

گیربکس نوع یک	۲۳	۲۳	۲۰	۲۱	
گیربکس نوع دو	۲۷	۲۹	۲۵	۲۳	۲۴
گیربکس نوع سه	۲۴	۲۶	۲۴		

در اینجا ملاحظه می‌شود که حجم نمونه‌ها برابر نیستند و داریم n_1=۴، n_2=۵ و n_3=۳. باید فرضیه‌ی $H_0 : \mu_1 = \mu_2 = \mu_3$ را در مقابل فرضیه‌ی مخالف عدم تساوی میانگین‌ها بیازماییم. در این مثال سه گروه داریم. بنابراین $k = 3$ و تعداد کل نمونه‌ها $N = 12$ می‌باشد. میانگین‌های هر کدام از گروه‌ها و میانگین کل به صورت زیرهستند:

$$\overline{X}_1 = 21.8\ ,\ \overline{X}_2 = 25.6\ ,\ \overline{X}_3 = 24.7\ ,\ \overline{X} = 24.1\ .$$

لذا واریانس بین گروه‌ها برابر است با

واریانس بین گروه‌ها= $MS_B = \frac{1}{3-1}\left[4(21.8-24.1)^2 + 5(25.6-24.1)^2 + 3(24.7-24.1)^2\right] = 16.7$

از طرفی داریم:

$$S_1^2 = \frac{1}{4-1}\left[(23-21.8)^2 + (23-21.8)^2 + (20-21.8)^2 + (21-21.8)^2\right] = 2.3$$
$$S_2^2 = \frac{1}{5-1}\left[(27-25.6)^2 + (29-25.6)^2 + (25-25.6)^2 + (23-25.6)^2 + (24-25.6)^2\right] = 5.8$$
$$S_3^2 = \frac{1}{3-1}\left[(24-24.7)^2 + (26-24.7)^2 + (24-24.7)^2\right] = 1.3$$

واریانس داخل گروه‌ها= $MS_w = \frac{1}{12-3}\left[(4-1)\times 2.3 + (5-1)\times 5.8 + (3-1)\times 1.3\right] = 3.6$

بنابراین، مقدار آماره‌ی آزمون به صورت $F = \frac{16.7}{3.6} = 4.74$ محاسبه می‌گردد.

با توجه به سطح معنی‌داری ۵ درصد و درجه‌های آزادی ۲ و ۹ از روی جدول توزیع فیشر مقدار بحرانی برابر ۴/۲۶ می‌باشد. لذا چون مقدار آماره‌ی آزمون از مقدار بحرانی بزرگ‌تر می‌باشد، می‌توان فرضیه‌ی اولیه مبنی بر تساوی میانگین‌ها را در سطح خطای ۵ درصد رد نمود به این معنی که به طور متوسط میزان مصرف بنزین اتومبیل‌ها با توجه به نوع گیربکس فرق می‌کند.

جدول آنالیز واریانس مثال فوق به صورت زیر است.

منبع تغییرات	مقدار	درجه آزادی	میانگین مربعات یا واریانس	آماره F
تبیین شده	۳۴/۳	۳-۱=۲	۱۷/۱۵	۴/۷۴
تبیین نشده	۳۲/۶	۱۲-۳=۹	۳/۶۲	--
جمع	۶۶/۹	۱۲-۱=۱۱	--	--

جلسه سوم

۱. بررسی پذیره‌های زیربنایی آنالیز واریانس یک طرفه
۲. آزمون‌های تعقیبی مقایسه‌ی چندگانه
۳. تخمین فاصله‌ای

همان‌طور که مشاهده نمودید، در مسایل تحلیل واریانس می‌خواهیم یک مدل به داده‌ها برازش داده و اثر عامل یا تیمار را بر متغیر پاسخ که همان متغیر وابسته با مقادیر کمی X است، بررسی کنیم. مدل آنالیز واریانس یک طرفه با اثر τ_i ($i=1,2,...,k$)، میانگین کل μ و خطاهای ε_{ij} به صورت زیر است:

$$X_{ij}=\mu+\tau_i+\varepsilon_{ij}, \qquad i=1,2,...,k,\ j=1,2,...,n_i$$

پس از این که مدل آنالیز واریانس برای مقایسه‌ی میانگین چند گروه مستقل انجام گرفت، مسأله‌ی مهمی که باید مورد بررسی قرار گیرد، پذیره‌هایی زیربنایی این مدل است.

پذیره‌های زیربنایی

پذیره‌های زیربنایی مدل آنالیز واریانس عبارتند از:

- میانگین خطاها باید صفر باشد
- خطاهای باید ناخودهمبسته (مستقل) باشند
- واریانس خطاها باید ثابت باشد
- خطاها باید هم‌توزیع و دارای توزیع نرمال باشند

تذکر

دو پذیره‌ی زیربنایی آخر از بقیه اهمیت بیشتری دارند.

بعد از انجام آنالیز واریانس یک طرفه لازم است که خطاهای مدل را یافته و با توجه به روش‌های نموداری یا آزمون فرضیه به بررسی پذیره‌های زیربنایی پرداخت. برای مدل فوق خطاهای نمونه مدل به صورت زیر محاسبه می‌شود:

$$e_{ij}=X_{ij}-\overline{X}_i, \qquad i=1,2,...,k,\ j=1,2,...,n_i$$

در رابطه‌ی فوق، e_{ij}ها برآورد خطاهای مدل در جامعه یعنی ε_{ij}ها می‌باشد.

با رسم نمودار توزیع فراوانی و نمودار احتمال نرمال یا احتمال تجمعی نرمال خطاهای نمونه، می‌توان در مورد پذیره‌ی زیربنایی بررسی توزیع خطاهای مدل، استنباط انجام داد. همچنین با استفاده از آزمون‌های

آماری کلموگروف-اسمیرنوف، شاپیرو-ویلک، اندرسون دارلینگ و غیره که در اکثر نرم‌افزارهای آماری موجود می‌باشند، می‌توان این پذیره را مورد آزمون قرار داد. لازم به ذکر است که در این گونه تحلیل‌ها، انحراف جزیی از نرمال بودن زیاد نگران کننده نیست.

برای بررسی پذیره‌ی زیربنایی ثابت بودن واریانس خطاهای مدل، باید نمودار خطاهای نمونه را نسبت به مقادیر برآورد شده از مدل ($\hat{X}_{ij}$) رسم نمود. مقادیر برآورد شده از مدل به صورت زیر محاسبه می‌گردند:

$$\hat{X}_{ij} = \hat{\mu} + \hat{\tau}_i = \overline{X} + \left(\overline{X}_i - \overline{X}\right) = \overline{X}_i$$

اگر در نمودار پراکنش خطاهای مدل در مقابل مقادیر برآورد شده، هیچ‌گونه روندی مبنی بر افزایش یا کاهش مقادیر مانده‌ها با توجه به تغییرات مقادیر برآورد شده مشاهده نگردد، می‌توان گفت که خطاهای مدل واریانس ثابت دارند.

بررسی پذیره‌های زیربنایی را در قسمت‌های بعدی با استفاده از نرم افزارهای آماری توزیع خواهیم داد.

آزمون‌های تعقیبی و مقایسه‌ی چندگانه

- **کم‌ترین تفاوت معنی‌دار**
- **توکی**
- **نیومن-کولز**
- **دانت**
- **دانکن**
- **شفه**
- **...**

پس از این که آزمون مقایسه‌ی میانگین چند جامعه‌ی مستقل را با استفاده از آنالیز واریانس انجام دادیم، می‌خواهیم ببینیم آیا میانگین‌های هر دو گروه با درنظر گرفتن اثرات گروه‌های دیگر با هم اختلاف معنی‌دار دارند یا خیر. به عبارت دیگر، می‌خواهیم آزمون مقایسه‌ی دو به دوی میانگین‌ها را انجام دهیم چون این آزمون‌ها بعد از آنالیز واریانس انجام می‌شوند، آزمون‌های تعقیبی آنالیز واریانس نیز نامیده می‌شوند.

در این آزمون‌های تعقیبی به دنبال آزمون فرضیه‌های $\mathrm{H}_0 : \mu_i = \mu_j , i \neq j , i = 1,2,..,k$ در مقابل فرضیه‌های جانشین هستیم. بعضی از این آزمون‌ها فقط در صورتی قابل انجام هستند که مدل آنالیز واریانس معنی‌دار باشد (یعنی فرضیه‌ی اولیه در آنالیز واریانس رد شود) و بعضی نیز چه مدل آنالیز واریانس معنی‌دار باشد یا نباشد، قابل استفاده هستند. شماری از انواع این آزمون‌های تعقیبی عبارتند از: آزمون کم‌ترین تفاوت معنی‌دار (LSD)، توکی، نیومن-کولز، دانت، دانکن، شفه و ....

تفاوت بین روش‌ها

نکته قابل توجه این است که نتایج بدست آمده از هر کدام از این روش‌ها، لزوماً یکسان نیستند، بدین معنی که ممکن است میانگین دو گروه با استفاده از یکی از آزمون‌ها اختلاف معنی‌دار داشته باشد و با استفاده از روش دیگر، اختلاف معنی‌داری نداشته باشد ولی نمی‌توان به صورت علمی نشان داد که کدام روش بهتر است. انتخاب نوع آزمون تعقیبی با توجه به تجارب و سلایق آماردان صورت می‌گیرد. اما بعضی از روش‌ها همانند دانکن و شفه نسبت به بقیه روش‌ها محافظه کارتر هستند و کوچک‌ترین اختلاف معنی‌دار را تشخیص می‌دهند و معمولاً در مسایل تجربی یا نیمه تجربی حساس، بیش‌تر مورد استفاده قرار می‌گیرند.

آزمون کم‌ترین تفاوت معنی‌دار

برای انجام این آزمون لازم است که ابتدا میانگین‌های گروه‌ها را با استفاده از آنالیز واریانس با همدیگر مقایسه نماییم. حال اگر فرضیه‌ی برابری میانگین‌های گروه‌ها رد شد (مدل آنالیز واریانس معنی‌دار باشد)، برای مقایسه‌ی دو به دوی میانگین گروه‌ها، می‌توان از آزمون کم‌ترین تفاوت معنی‌دار استفاده نمود. از این‌رو برای دو گروه i و j ($i, j = 1,2,\dots,k, \quad i \neq j$) در صورتی می‌توان گفت میانگین‌های دو گروه اختلاف معنی‌دار دارند که:

$$\left|\overline{X}_i - \overline{X}_j\right| \geq t_{1-\frac{\alpha}{2},N-k}\sqrt{MS_w\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

که در آن MS_w میانگین مربعات خطا (داخل گروه‌ها)، N تعداد کل مشاهدات، k تعداد گروه‌ها و $t_{1-\frac{\alpha}{2},N-k}$ صدک $1-\frac{\alpha}{2}$ توزیع استودنت با درجه‌ی آزادی $N-k$ است.

مثال ۶. سه نوع گیربکس مختلف روی اتومبیل‌های مختلفی که همگی موتورهای شش سیلندر با حجم موتور ۳۰۰۰ میلی لیتر دارند، کار گذاشته شده است. مصرف بنزین به مایل بر گالن نتایج زیر را دربر داشته است. در سطح معنی‌داری پنج درصد، این ادعا را آزمون کنید که متوسط میزان مصرف بنزین خودروهای با گیربکس‌های مختلف، متفاوت نمی‌باشد. در سطح خطای ۵ درصد و در صورت بروز اختلاف معنی‌دار، با استفاده از آزمون کم‌ترین تفاوت معنی‌دار مشخص کنید کدام دو گروه باعث این اختلاف شده است؟

گیربکس نوع یک	۲۳	۲۳	۲۰	۲۱	
گیربکس نوع دو	۲۷	۲۹	۲۵	۲۳	۲۴
گیربکس نوع سه	۲۴	۲۶	۲۴		

در جلسه‌ی قبل به بررسی ادعا با استفاده از آنالیز واریانس پرداخته شد. در آن‌جا دیدیم که در سطح خطای ۵ درصد، متوسط میزان مصرف بنزین خودروها با سه نوع گیربکس متفاوت است. حال برای بررسی این

که کدام یک از سه گیربکس باعث اختلاف معنی‌دار شده از آزمون تعقیبی کم‌ترین تفاوت معنی‌دار به نحو زیر استفاده می‌کنیم.

چون $MS_w = 3.6$ ، $N-k=9$ و $t_{0.975,9} = 2.26$، لذا

$$\bar{X}_1 = 21.8\ ,\ \bar{X}_2 = 25.6\ ,\ \bar{X}_3 = 24.7\ ,$$

$$\left|\bar{X}_1 - \bar{X}_2\right| = |۲۱/۸-۲۵/۶| = ۳/۸, \qquad t_{1-\frac{\alpha}{2},N-k}\sqrt{MS_w\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = 2.26\sqrt{3.6\left(\tfrac{1}{4}+\tfrac{1}{5}\right)} = 2.88$$

$$\left|\bar{X}_1 - \bar{X}_3\right| = |۲۱/۸-۲۴/۷| = ۲/۹, \qquad t_{1-\frac{\alpha}{2},N-k}\sqrt{MS_w\left(\frac{1}{n_1}+\frac{1}{n_3}\right)} = 2.26\sqrt{3.6\left(\tfrac{1}{4}+\tfrac{1}{3}\right)} = 3.28$$

$$\left|\bar{X}_2 - \bar{X}_3\right| = |۲۵/۶-۲۴/۷| = ۰/۹, \qquad t_{1-\frac{\alpha}{2},N-k}\sqrt{MS_w\left(\frac{1}{n_3}+\frac{1}{n_3}\right)} = 2.26\sqrt{3.6\left(\tfrac{1}{5}+\tfrac{1}{3}\right)} = 3.13$$

لذا با توجه به مقادیر بدست آمده و مقایسه آن با $t_{1-\frac{\alpha}{2},N-k}\sqrt{MS_w\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$ می‌توان دریافت که میانگین دو گروه اول و دوم با هم اختلاف معنی‌دار در سطح ۵ درصد داشته اند اما میانگین دو گروه اول و سوم و دو گروه دوم و سوم اختلاف معنی‌داری با توجه به آزمون کمترین تفاوت معنی‌دار نداشته اند.

آزمون چنددامنه‌ای دانکن

آزمون چند دامنه‌ای دانکن یکی از متداول‌ترین آزمون‌های تعقیبی برای مقایسه‌ی تمام جفت‌های میانگین‌ها است. بر خلاف آزمون‌های LSD و دانت که تنها پس از معنی‌دار بودن مدل قابل استفاده‌اند، آزمون دانکن را همیشه و حتی وقتی مدل تحلیل واریانس معنی‌دار نباشد، می‌توان بکار برد. در حالتی که میانگین‌ها به یکدیگر نزدیک‌اند و در طرفین میانگین کل طوری قرار دارند که اختلاف آن‌ها نسبت به میانگین کل کم است، ممکن است این اختلاف‌ها یکدیگر را خنثی کرده و مدل معنی‌دار نشود. این درحالی است که ممکن است بین کوچک‌ترین و بزرگ‌ترین میانگین‌ها اختلاف معنی‌داری وجود داشته باشد. در آزمون دانکن سطوح متغیر عامل دو به دو با یکدیگر مقایسه می‌شوند و نیازی به وجود شاهد نیست. برای انجام این آزمون به ترتیب زیر عمل می‌کنیم:

۱. میانگین‌ها را به ترتیب افزایشی مرتب نموده و خطای معیار میانگین‌ها را از فرمول $S_{\bar{X}_i} = \sqrt{\frac{MS_w}{n_h}}$ حساب می‌کنیم که در آن n_h میانگین هارمونیک (هم‌ساز) تعداد تکرارها (نمونه‌های تصادفی از هر گروه) است ($n_h = \frac{k}{\sum_{i=1}^{k}\frac{1}{n_i}}$). توجه کنید وقتی حجم نمونه‌های انتخابی از گروه‌ها مساوی و برابر n باشد، $n_h = n$ است.

۲. از جدول دامنه‌های معنی‌دار دانکن که معمولاً در اکثر کتب مربوط به طرح آزمایش‌های آماری وجود دارد، برای $p=2,3,\dots,k$ مقادیر $r_\alpha(p,f)$ را که در آن α سطح معنی‌داری آزمون و f درجه‌ی آزادی خطا است، بدست می‌آوریم.

۳. با محاسبه $R_p=r_\alpha(p,f)S_{\bar{X}_i}\,,\quad p=2,3,\dots,k$
مجموعه‌ای شامل $k-1$ مقدار را تحت عنوان کم‌ترین دامنه‌های معنی‌دار به دست می‌آوریم.

۴. اختلاف میانگین‌ها را تک تک با شروع از بزرگ‌ترین نسبت به کوچک‌ترین آن‌ها که با R_k مقایسه می‌شود، آزمون می‌کنیم.

فرض کنید ردیف میانگین‌ها را به ترتیب نزولی ۱، ۲،...، k قرار دهیم. ابتدا اختلاف میانگین مرتبه اول از میانگین مرتبه k را با R_k مقایسه می‌کنیم. اگر این اختلاف بیش‌تر از R_k بود، نتیجه می‌گیریم که جفت میانگین‌های اول و آخر به صورتی معنی‌دار متفاوت‌اند. در غیر این صورت، یعنی وقتی اختلاف میانگین‌های مرتبه اول و آخر از R_k کمتر است، می‌گوییم بزرگ‌ترین و کوچک‌ترین میانگین‌ها تفاوتی با هم ندارند. در چنین حالتی اختلاف هیچ جفت میانگینی را معنی‌دار نمی‌دانیم. پس از آن اختلاف میانگین مرتبه اول و میانگین مرتبه $k-1$ را با R_{k-1} مقایسه می‌کنیم. اختلاف میانگین مرتبه اول از میانگین مرتبه $k-2$ و بالاخره اختلاف میانگین مرتبه اول از میانگین مرتبه دوم را با R_2 مقایسه می‌کنیم. اختلاف میانگین مرتبه دوم از میانگین مرتبه kام با $R_{k-2+1}=R_{k-1}$ و به طور کلی اختلاف میانگین مرتبه j ام را نسبت به میانگین مرتبه iام با R_{i-j+1} مقایسه می‌کنیم.

همان‌طور که بیان شد، اگر اختلاف از کم‌ترین دامنه‌ی معنی‌دار متناظر بیش‌تر بود، آن را معنی‌دار گوییم. در غیر این صورت، اختلاف این دو و اختلاف تمام میانگین‌هایی که بین این دو قرار دارند، معنی‌دار نیستند.

مثال ۷. پژوهش‌گری می‌خواهد میزان باروری بذر گندم در اثر استفاده از کود نیترات آمونیوم را مطالعه نماید. در این آزمایش که به صورت کاملاً تصادفی اجرا می‌شود، پنج سطح کود درنظر گرفته شده و برای هر سطح نیز پنج تکرار وجود دارد. داده‌ها به شرح جدول ذیل است.

درصد کود/ تکرارها	۱	۲	۳	۴	۵	میانگین
۰	۷	۷	۱۵	۱۱	۹	۹/۸
۴	۱۲	۱۷	۱۲	۱۸	۱۸	۱۵/۴
۶	۱۴	۱۸	۱۸	۱۹	۱۹	۱۷/۶
۸	۱۹	۲۵	۲۲	۱۹	۲۳	۲۱/۶
۱۰	۷	۱۰	۱۱	۱۵	۱۱	۱۰/۸
میانگین						۱۵/۰۴

پس از محاسبات می‌دانیم که MS_w=۸/۰۶ و $n_h = n$=۵ و میانگین گروه‌ها به ترتیب افزایشی به صورت زیر است:

$$\bar{X}_1 = 9.8, \bar{X}_5 = 10.8, \bar{X}_2 = 15.4, \bar{X}_3 = 17.6, \bar{X}_4 = 21.6,$$

و خطای معیار میانگین‌ها برابر است با

$$S_{\bar{X}_i} = \sqrt{\frac{MS_w}{n}} = \sqrt{\frac{8.06}{5}} = 1.27$$

از جدول دامنه‌های دانکن برای k=۵، p=۴،۳،۲ و f=۲۰ درجه آزادی برای خطا و سطح معنی‌داری پنج درصد بدست می‌آوریم:

$$r_{0.05}(2,20) = 2.95\,,\ r_{0.05}(3,20) = 3.10\,,\ r_{0.05}(4,20) = 3.18\,,\ r_{0.05}(5,20) = 3.25$$

از ضرب این مقادیر در خطای معیار میانگین‌ها، کمترین دامنه‌های معنی‌داربه صورت زیر است:

$$R_2 = r_{0.05}(2,20)S_{\bar{X}_i} = 3.75\,,\ R_3 = r_{0.05}(3,20)S_{\bar{X}_i} = 3.94\,,$$
$$R_4 = r_{0.05}(4,20)S_{\bar{X}_i} = 4.04\,,\ R_5 = r_{0.05}(5,20)S_{\bar{X}_i} = 4.13$$

حال اختلاف میانگین‌ها را با شروع از بزرگ‌ترین نسبت به کوچک‌ترین محاسبه و آن‌ها را به شرح زیر با کم‌ترین دامنه‌های معنی‌دار مقایسه می‌کنیم.

اختلاف میانگین گروه‌ها

۴ از گروه ۱: ۲۱/۶ −۹/۸=۱۱/۸> R_5 =۴/۱۳

۴ از گروه ۵: ۲۱/۶ −۱۰/۸=۱۰/۸> R_4 =۴/۰۴

۴ از گروه ۲: ۲۱/۶ −۱۵/۴=۶/۲> R_3 =۳/۹۴

۴ از گروه ۳: ۲۱/۶ −۱۷/۶=۴/۰> R_2 =۳/۷۵

۳ از گروه ۱: ۱۷/۶ −۹/۸=۷/۸> R_4 =۴/۰۴

۳ از گروه ۵: ۱۷/۶ −۱۰/۸=۶/۸> R_3 =۳/۹۴

۳ از گروه ۲: ۱۷/۶ −۱۵/۴=۲/۲< R_2 =۳/۷۵

۲ از گروه ۱: ۱۵/۴ −۹/۸=۵/۶> R_3 =۳/۹۴

۲ از گروه ۵: ۱۵/۴ −۱۰/۸=۴/۶> R_2 =۳/۷۵

۵ از گروه ۱: ۱۰/۸ −۹/۸=۱/۰< R_2 =۳/۷۵

پس تمام میانگین‌ها را به جز میانگین‌های مرتبه دوم و سوم و میانگین‌های مرتبه پنجم و اول با هم اختلاف معنی‌دار دارند.

بازه‌های اطمینان

پس از تحلیل واریانس معمولاً علاقه‌مند به تعیین بازه‌های اطمینان هستیم. واضح است که در این مدل تحلیل واریانس $\overline{X}_i$ ها متغیرهای تصادفی مستقل و دارای توزیع نرمال با میانگین μ_i و واریانس σ^2/n هستند. اگر σ^2 (واریانس جامعه) معلوم باشد، می‌توانیم برای تعیین بازه‌ی اطمینان از توزیع نرمال استفاده کنیم. اما اگر σ^2 نامعلوم باشد، می‌توانیم با جایگذاری میانگین مربعات خطا (داخل گروه‌ها) به جای σ^2 بازه اطمینان را براساس توزیع استودنت بسازیم. اگر σ^2 مجهول باشد، بازه‌ی اطمینان $(1-\alpha)$ ۱۰۰ درصد برای μ_i، یعنی **i**امین میانگین تیماری عبارت است از:

$$\overline{X}_i \pm t_{1-\frac{\alpha}{2},N-k}\sqrt{\frac{MS_w}{n_i}}$$

بازه اطمینان $(1-\alpha)$ ۱۰۰ درصد برای تفاوت دو میانگین تیماری مثلاً $\mu_i - \mu_j$ نیز

$$\overline{X}_i - \overline{X}_j \pm t_{1-\frac{\alpha}{2},N-k}\sqrt{MS_w(\frac{1}{n_i}+\frac{1}{n_j})}$$

است.

مثال ۸. در مثال مربوط به میزان باروری بذر گندم در اثر استفاده از کود نیترات آمونیوم، می‌خواهیم یک بازه اطمینان ۹۵ درصد برای میانگین تیمار چهارم و یک فاصله اطمینان ۹۵ درصد برای تفاوت دو میانگین تیمارهای پنج و یک تعیین کنیم. داریم:

$$\overline{X}_i \pm t_{1-\frac{\alpha}{2},N-k}\sqrt{\frac{MS_w}{n}} = 21.6 \pm 2.086\sqrt{\frac{8.06}{5}} = 21.6 \pm 2.65$$

بنابراین بازه اطمینان ۹۵ درصد برای میانگین گروه چهارم عبارت است از (۲۴/۲۵ و ۱۸/۹۵). همچنین

$$\overline{X}_i - \overline{X}_i \pm t_{1-\frac{\alpha}{2},N-k}\sqrt{\frac{2MS_w}{n}} = 10.8 - 9.8 \pm 2.086\sqrt{\frac{2\times 8.06}{5}} = 1 \pm 3.75$$

و بازه اطمینان ۹۵ درصد برای تفاوت میانگین‌های دو گروه پنج و یک به صورت (۴/۷۵ و ۲/۷۵-)، می‌باشد.

جلسه چهارم

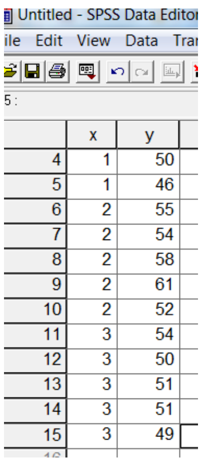
۱. استفاده از نرم افزارهای آماری

۲. تفسیر خروجی مدل آنالیز واریانس و آزمون‌های تعقیبی

مثال ۱ با نرم افزار SPSS

چون در اینجا به مقایسه‌ی گروه‌های مستقل می‌پردازیم، باید در صفحه‌ی داده‌های نرم‌افزار (Data view) یک متغیر (ستون X) تحت عنوان متغیر مستقل که نشان‌دهنده‌ی سه گروه یا سه نوع ماشین باشد، تعریف کنیم و یک متغیر (ستون Y) دیگر تحت عنوان متغیر وابسته که در واقع مشاهدات مربوط به میزان محصولات را در سه گروه نشان می‌دهد، ایجاد نماییم (شکل زیر).

سپس با استفاده از منوی Analyze>Compare Means>Oneway ANOVA می‌توانیم مقایسه‌ی مزبور را انجام دهیم. خروجی این مثال به صورت زیر است.



	x	y
4	1	50
5	1	46
6	2	55
7	2	54
8	2	58
9	2	61
10	2	52
11	3	54
12	3	50
13	3	51
14	3	51
15	3	49

ANOVA

Product

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	130.000	2	65.000	8.298	.005
Within Groups	94.000	12	7.833		
Total	224.000	14			

جدول فوق همان جدول آنالیز واریانس یک طرفه می‌باشد که شاخص‌های تعریف شده در قسمت مربوطه را در خود جای داده است. چون مقدار آماره‌ی آزمون بزرگ‌تر از مقدار ناحیه‌ی بحرانی می‌باشد یا این که p–مقدار (ستون .Sig) آزمون کمتر از ۵ درصد است، فرضیه‌ی صفر مبنی بر برابری میانگین میزان محصولات سه ماشین در سطح خطای ۵ درصد رد می‌شود.

مثال ۲ با نرم افزار SPSS

همانند مثال یک داده‌ها را در نرم افزار وارد نموده و داریم:

ANOVA

Product

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.133	2	6.067	.879	.440
Within Groups	82.800	12	6.900		
Total	94.933	14			

در اینجا چون آماره‌ی آزمون از مقدار ناحیه‌ی بحرانی کم‌تر است یا این که p-مقدار آزمون از ۵ درصد بیش‌تر می‌باشد، فرضیه صفر یعنی برابری میانگین محصولات سه گروه در سطح خطای پنج درصد رد نمی‌شود.

مثال ۳ با نرم افزار SPSS

همانند قبل داده‌ها را وارد نموده و مورد تجزیه و تحلیل قرار می‌دهیم.

ANOVA

Product

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	130.000	2	65.000	1.130	.355
Within Groups	690.000	12	57.500		
Total	820.000	14			

چون آماره‌ی آزمون از مقدار ناحیه‌ی بحرانی کم‌تر است و یا p-مقدار بیش‌تر از پنج درصد می‌باشد، فرضیه اولیه در سطح خطای پنج درصد رد نمی‌شود.

مثال ۴ با نرم افزار SPSS

همانند قبل داده‌ها را وارد نموده و مورد تجزیه و تحلیل قرار می‌دهیم. خروجی جدول آنالیز واریانس به شکل زیر است.

ANOVA

Lifetime

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.468	2	2.234	1.353	.295
Within Groups	19.816	12	1.651		
Total	24.284	14			

آماره‌ی آزمون از مقدار ناحیه بحرانی کم‌تر است و همچنین p-مقدار بیش‌تر از پنج درصد می‌باشد، لذا فرضیه‌ی اولیه در سطح خطای پنج درصد رد نمی‌شود.

مثال ۵ با نرم افزار SPSS

همانند قبل داده‌ها را وارد نموده و مورد تجزیه و تحلیل قرار می‌دهیم.

ANOVA

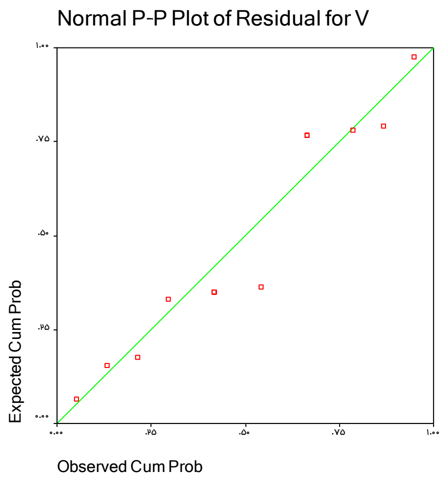
Oil

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	34.300	2	17.150	4.732	.039
Within Groups	32.617	9	3.624		
Total	66.917	11			

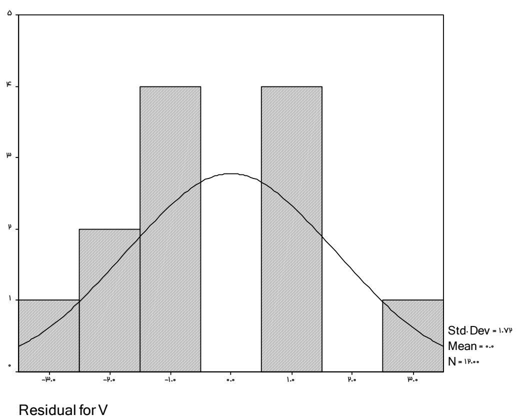
آماره‌ی آزمون از مقدار ناحیه‌ی بحرانی بیش‌تر است و یا این که مشاهده می‌کنیم p-مقدار کم‌تر از پنج درصد می‌باشد، لذا فرضیه‌ی اولیه یعنی تساوی مصرف بنزین خودورها با سه نوع گیربکس در سطح خطای پنج درصد رد می‌شود.

در ادامه برای همین مثال می‌خواهیم پذیره‌های زیربنایی مدل آنالیز واریانس یک طرفه را بررسی نماییم. برای این منظور باید نمودارهای توزیع فراوانی، احتمال نرمال و احتمال تجمعی نرمال را برای خطاهای مدل جهت بررسی توزیع خطاها و نمودار پراکنش مقادیر برآورد شده از مدل در مقابل خطاها را برای بررسی پذیره ثابت بودن واریانس خطاها، رسم کنیم و همان‌طور که بیان شد، با توجه به نمودارها تصمیم‌گیری نماییم. بدیهی است که پذیره‌ی زیربنایی استقلال خطاها نیز مستقیماً با انتخاب تصادفی مشاهدات از جامعه مرتبط است. بنابراین در صورت انتخاب تصادفی مشاهدات، این پذیره برقرار خواهد بود. برای رسم نمودارها باید از منوی Analyze>General Linear Model>Univariate استفاده نموده و متغیر مستقل یا گروه را در قسمت Fixed Factor(s) وارد نموده و متغیر وابسته را به قسمت Dependent Variable انتقال داد. سپس از گزینه Save در همین پنجره استفاده نموده و گزینه‌های Unstandardized را برای Residuals (خطاها) و Predicted Values (مقادیر برآورد شده) را علامت زد. حال این دو متغیر به جدول متغیرها اضافه خواهد شد که برای رسم نمودار توزیع فراوانی از Graphs>Histogram، برای رسم نمودار احتمال نرمال از Graphs>P-P، نمودار احتمال تجمعی نرمال از Graphs>Q-Q استفاده نموده و متغیر موردنظر در این دستورات را خطاهای ذخیره شده در انتهای ستون داده‌ها درنظر می‌گیریم. سپس نمودار پراکنش مقادیر برآورد شده در مقابل خطاها را از منوی Graphs>Scatter رسم می‌کنیم. همان‌طور که قبلاً تأکید شد، برای بررسی پذیره‌های زیربنایی می‌توان از روش‌های آزمون فرضیه نیز استفاده کرد. در این صورت برای بررسی پذیره نرمال بودن توزیع خطاهای مدل می‌توان از آزمون کلموگروف-اسمیرنوف استفاده نمود. این آزمون از منوی Analyze>Nonparametric Tests>1-Sampe K-S قابل انجام می‌باشد. برای بررسی ثابت بودن خطاها نیز می‌توان در دستور Analyze>General Linear Model>Univariate و قسمت

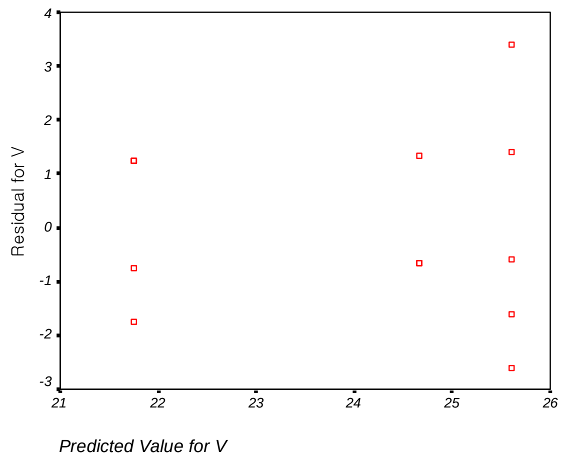
Options، گزینه‌ی Homogeneity tests را انتخاب نموده تا نتیجه‌ی آزمون همگن بودن داده‌ها به خروجی نرم‌افزار فرستاده شود. شکل‌های زیر به ترتیب هیستوگرام، نمودار احتمال نرمال، نمودار احتمال تجمعی نرمال و نمودار پراکنش خطاها در مقابل مقادیر برآورد شده را برای خطاهای نمونه‌ی مثال ۵ نشان می‌دهد.



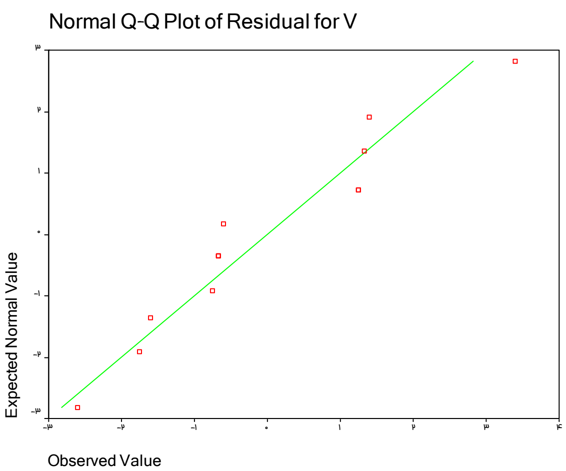
نمودار احتمال نرمال خطاها



هیستوگرام خطاها



نمودار پراکنش خطاها در مقابل مقادیر برآورد شده



نمودار احتمال تجمعی نرمال

چون نقاط روی نمودارهای احتمال و احتمال تجمعی نرمال تقریباً از ی خط راست پیروی می‌کنند، نتیجه می‌گیریم که توزیع خطاها تقریباً نرمال است. واریانس خطاها نیز به علت این که روند خاصی در نمودار پراکنش مقادیر خطاها در مقابل مقادیر برآورد شده وجود ندارد، ثابت است.

جداول زیر نتایج آزمون همگن بودن واریانس و آزمون کلموگروف-اسمیرنوف را نشان می‌دهد. با توجه به آزمون همگن بودن واریانس خطاها، چون p-مقدار کم‌تر از ۵ درصد نیست، نتیجه می‌گیریم که واریانس خطاها ثابت است. آزمون کلموگروف-اسمیرنوف نیز در سطح خطای ۵ درصد نشان می‌دهد که توزیع خطاها نرمال است (p=۰/۶۰۹-مقدار).

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Oil

F	df1	df2	Sig.
1.785	2	9	.222

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+T

آزمون همگن بودن واریانس خطاها

One-Sample Kolmogorov-Smirnov Test

		Residual for V
N		12
Normal Parameters [a,b]	Mean	.0000
	Std. Deviation	1.72196
Most Extreme Differences	Absolute	.220
	Positive	.220
	Negative	-.183
Kolmogorov-Smirnov Z		.761
Asymp. Sig. (2-tailed)		.609

a. Test distribution is Normal.

b. Calculated from data.

آزمون کلموگروف-اسمیرنوف

مثال ۶ با نرم افزار SPSS

همانند قبل داده‌ها را وارد نموده و مورد تجزیه و تحلیل قرار می‌دهیم. برای انجام آزمون کم‌ترین تفاوت معنی‌دار، باید آنالیز واریانس یک طرفه را از منوی مربوطه فراخوانده و قسمت Post HOC را انتخاب نمود. پس از باز شدن پنجره‌ی مربوطه، گزینه‌ی LSD را علامت می‌زنیم. خروجی این گزینه به صورت زیر خواهد بود.

Multiple Comparisons

Dependent Variable: Oil
LSD

(I) Gear Box Type	(J) Gear Box Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	2	-3.8500*	1.27704	.015	-6.7389	-.9611
	3	-2.9167	1.45397	.076	-6.2058	.3725
2	1	3.8500*	1.27704	.015	.9611	6.7389
	3	.9333	1.39027	.519	-2.2117	4.0783
3	1	2.9167	1.45397	.076	-.3725	6.2058
	2	-.9333	1.39027	.519	-4.0783	2.2117

*. The mean difference is significant at the .05 level.

همان‌طور که مشاهده می‌شود، در سطح خطای پنج درصد فقط گروه‌های اول و دوم اختلاف معنی‌دار دارند (آن‌هایی که علامت ستاره دارند).

مثال ۷ با نرم افزار SPSS

همانند قبل داده‌ها را ثبت کرده و مورد تجزیه و تحلیل قرار می‌دهیم. برای انجام آزمون دانکن، باید آنالیز واریانس یک طرفه را از منوی مربوطه فراخوانده و در قسمت Post HOC گزینه‌ی Duncan را علامت بزنیم. خروجی آزمون دانکن به صورت جدول زیر است.

Wheat Product

Duncan[a,b]

Amunium Nitrat Level	N	Subset 1	2	3
0	5	9.8000		
10	5	10.8000		
4	5		15.4000	
6	5		17.6000	
8	5			21.6000
Sig.		.584	.235	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 8.060.

a. Uses Harmonic Mean Sample Size = 5.000.

b. Alpha = .05.

همان‌طور که در جدول فوق دیده می‌شود با توجه به آزمون دانکن در سطح خطای ۵ درصد، فقط سطوح کود صفر و ده درصد و چهار و شش درصد اختلاف معنی‌داری نداشته و بقیه دارای تفاوت معنی‌دار هستند (میانگین‌هایی که در دو ستون جدا از هم قرار می‌گیرند). بنابراین تفاوت‌های معنی‌دار عبارتند از:

- سطح کود صفر درصد با تیمارهای ۴، ۶ و ۸ درصد
- سطح کود ۱۰ با تیمارهای ۴، ۶ و ۸ درصد
- سطح کود ۴ درصد با کود ۸ درصد
- کود ۶ درصد با کود ۸ درصد