



صفر تا صد آمار زیستی

مؤلف:

لیلا تیموری یگانه



موسسه علمی انتشاراتی سنا
(سامانه نوین آموز)



آمار زیستی شاخه‌ای از آمار کاربردی است که تمرکز و تأکید آن بر توسعه و استفاده از روش‌های آماری است که در حل مسائل و پاسخ به سؤالات حوزه بهداشت، پزشکی، ژنتیک و بیولوژی انسانی مورد استفاده قرار می‌گیرند. برای مثال خبر کشف رابطه بین استعمال سیگار و یک بیماری جدید حاصل استفاده از علم آمار در تحقیقات پزشکی است. به‌طوری‌که بدون فهم و آشنایی با مباحث آمار زیستی، نمی‌توان روش‌های جدید و مناسبی برای یافتن ژن‌های مستعد بیماری‌های مختلف و نحوه ارتباط آنها با یکدیگر یافت با توجه به گستردگی و اهمیت این درس نحوه و نوع آموزش آن برای فراگیری بهتر مخاطبان نیز حائز اهمیت است. با توجه به اینکه آمار زیستی یکی از دروس مهم آمار برای ورود به تحصیلات تکمیلی و دروس دانشگاه در مقطع کارشناسی می‌باشد و مخاطبین این رشته نیز کلیه کارشناسی‌های علوم پزشکی، کارشناس آمار، ریاضی، علوم کامپیوتر، دکتری عمومی گروه پزشکی و دکتری حرفه‌ای دامپزشکی می‌باشد. سرفصل‌های این درس به‌گونه‌ای تدوین شده است که اکثراً نیازهای دانشجویان را برآورده نماید. امید است کتاب حاضر مورد توجه شما عزیزان قرار گرفته باشد.

لیلا تیموری یگانه

کارشناس ارشد آمار گرایش ریاضی

عنوان	صفحه
پیشگفتار.....	۴

فصل اول: آمار توصیفی

مقدمه.....	۱۳
جامعه آماری.....	۱۳
نمونه.....	۱۴
انواع متغیر.....	۱۵
الف) طبقه‌بندی متغیر برحسب ماهیت متغیر.....	۱۵
ب) طبقه‌بندی متغیرها برحسب نقش متغیر در تحقیق.....	۱۶
رابطه بین متغیرها.....	۲۰
جدول توزیع فراوانی دسته‌بندی نشده.....	۲۰
جدول توزیع فراوانی دسته‌بندی شده.....	۲۱
قواعد کلی برای تشکیل جدول توزیع فراوانی طبقه‌بندی شده.....	۲۲
محاسبه دامنه تغییرات.....	۲۳
محاسبه فاصله طبقات.....	۲۳
محاسبه انواع فراوانی.....	۲۵
محاسبه نماینده طبقه	۲۵
محاسبه حدود واقعی طبقه.....	۲۶
توصیف داده به وسیله نمودارها.....	۲۶
معیارهای گرایش به مرکز.....	۲۷
رابطه بین اندازه‌های گرایش به مرکز.....	۳۲
دامنه تغییرات.....	۳۲
مجموعه سؤالات کنکوری آمار توصیفی.....	۳۵
پاسخ مجموعه سؤالات کنکوری آمار توصیفی.....	۳۷

فصل دوم: آنالیز ترکیبی و احتمال

مقدمه.....	۳۹
تعمیم اصل ضرب	۳۹
جایگشت‌های n تایی n کاراکتر متمایز	۴۰
جایگشت‌های r تایی n کاراکتر متمایز	۴۰

.....	جایگشت‌های دور میزگرد	۴۰
.....	جایگشت تکراری	۴۰
.....	مدل توپ و ظرف	۴۱
.....	محاسبه احتمال و اعمال جبری روی پیشامدها	۴۳
.....	انواع احتمال	۴۴
.....	احتمال شرطی، فرمول بیز، استقلال پیشامدها	۴۵
.....	قضیه احتمال کل	۴۵
.....	استقلال پیشامدها	۴۶
.....	مسائل ماکزیم و مینیم	۴۶
.....	احتمال شرطی	۴۷
.....	قانون ضرب احتمال	۴۸
.....	قضیه بیز	۴۸
.....	مجموعه سؤالات کنکوری آنالیز ترکیبی و احتمال	۴۹
.....	پاسخ مجموع سؤالات کنکوری آنالیز ترکیبی و احتمال	۵۴

فصل سوم: انواع متغیرهای تصادفی و ضریب همبستگی

.....	مقدمه	۵۹
.....	تابع توزیع احتمال	۵۹
.....	تابع توزیع تجمعی	۵۹
.....	انواع متغیرهای تصادفی	۶۰
.....	امید ریاضی	۶۳
.....	واریانس (پراکندگی)	۶۵
.....	برآورد ضریب همبستگی	۶۶
.....	ویژگی ضریب همبستگی	۶۶
.....	توزیع تقریبی R	۶۷
.....	مجموعه سؤالات کنکوری انواع متغیرهای تصادفی و ضریب همبستگی	۶۹
.....	پاسخ مجموعه سؤالات کنکوری انواع متغیرهای تصادفی و ضریب همبستگی	۸۰

فصل چهارم: توزیع‌های خاص

.....	مقدمه	۸۹
.....	توزیع دو جمله‌ای	۸۹
.....	توزیع پواسون	۹۰
.....	توزیع هندسی	۹۱

توزیع دوجمله‌ای منفی.....	۹۱
متغیر تصادفی فوق هندسی.....	۹۲
متغیر تصادفی یکنواخت گسسته.....	۹۳
متغیر تصادفی پیوسته یکنواخت.....	۹۴
متغیر تصادفی نرمال یا گاوسی.....	۹۵
متغیر تصادفی نمایی.....	۹۶
متغیر تصادفی گاما.....	۹۶
متغیر تصادفی بتا.....	۹۷
مجموعه سؤالات کنکوری توزیع‌های خاص	۹۸
پاسخ مجموعه سؤالات کنکوری توزیع‌های خاص	۱۰۷

فصل پنجم: فاصله اطمینان

مقدمه.....	۱۱۵
فاصله اطمینان میانگین توزیع نرمال برای جوامع با واریانس معلوم.....	۱۱۵
فاصله اطمینان میانگین توزیع نرمال برای جوامع با واریانس نامعلوم.....	۱۱۶
فاصله اطمینان واریانس توزیع نرمال.....	۱۱۸
فاصله اطمینان تفاوت میانگین دو جامعه نرمال.....	۱۱۸
فاصله اطمینان تفاضل میانگین دو جامعه نرمال با انحراف معیار نامعلوم و برابر.....	۱۲۰
نارایی برآوردگر.....	۱۲۲
میانگین مربعات خطا.....	۱۲۲
قضیه نامساوی کرامر راثو.....	۱۲۳
سازگاری برآوردگر.....	۱۲۳
حداکثر درستی‌مایی (MLE).....	۱۲۴
روش گشتاورها.....	۱۲۵
مجموعه سؤالات کنکوری فاصله اطمینان.....	۱۲۶
پاسخ مجموعه سؤالات کنکوری فاصله اطمینان.....	۱۳۶

فصل ششم: آزمون فرض

مقدمه.....	۱۴۵
تعاریف.....	۱۴۵
آزمون فرض میانگین جامعه نرمال با انحراف معیار معلوم.....	۱۴۶
آزمون فرض میانگین جامعه نرمال با انحراف معیار نامعلوم.....	۱۴۸

آزمون فرض در مورد واریانس یک جامعه نرمال.....	۱۴۹
آزمون فرض برابری میانگین دو جامعه نرمال مستقل، انحراف معیار جوامع معلوم.....	۱۵۰
آزمون فرض برابری میانگین دو جامعه نرمال مستقل، با انحراف معیار برابر و نامعلوم.....	۱۵۲
آزمون فرض برابری میانگین در مشاهدات زوجی.....	۱۵۳
آزمون فرض برابری واریانس دو جامعه نرمال.....	۱۵۳
مجموعه سؤالات کنکوری آزمون فرض.....	۱۵۵
پاسخ مجموعه سؤالات کنکوری آزمون فرض.....	۱۵۹

فصل هفتم: تحلیل واریانس و کوواریانس، آزمون‌های تعقیبی و طرح‌های بلوکی

تصادفی شده

آنالیز واریانس یک‌طرفه	۱۶۵
تحلیل واریانس بلوکی.....	۱۶۹
آزمون توکی یا HSD	۱۷۱
آزمون شفه (Scheffe test).....	۱۷۲
آزمون نیومن-کولز (The Newman-Keuls test)	۱۷۲
آزمون حداقل تفاوت معنی‌دار فیشر (LSD).....	۱۷۲
آزمون دانکن.....	۱۷۲
مجموعه سؤالات کنکوری تحلیل واریانس و کوواریانس، آزمون‌های تعقیبی و طرح‌های بلوکی تصادفی شده	۱۷۳
پاسخ مجموعه سؤالات کنکوری تحلیل واریانس و کوواریانس، آزمون‌های تعقیبی و طرح‌های بلوکی تصادفی شده	۱۸۴

فصل هشتم: نمونه‌گیری

مقدمه.....	۱۸۹
تعریف.....	۱۸۹
دلایل نمونه‌برداری.....	۱۸۹
نمونه‌برداری احتمالی.....	۱۸۹
نمونه‌گیری تصادفی.....	۱۹۰
نمونه‌گیری تصادفی با جایگذاری.....	۱۹۰
نمونه‌گیری تصادفی بدون جایگذاری.....	۱۹۰
طرح نمونه‌برداری نظام‌دار (سیستماتیک).....	۱۹۱
طرح نمونه‌برداری تصادفی طبقه‌ای.....	۱۹۲

نمونه برداری خوشه‌ای.....	۱۹۳
نمونه برداری سهمیه‌ای.....	۱۹۴
نمونه برداری مضاعف.....	۱۹۴
آماره.....	۱۹۵
توزیع مربع کای.....	۱۹۵
توزیع تی.....	۱۹۸
توزیع F.....	۱۹۹
توزیع میانگین نمونه برای جامعه متناهی.....	۲۰۰
توزیع میانگین نمونه برای جامعه نامتناهی.....	۲۰۱
توزیع واریانس نمونه برای جامعه نامتناهی.....	۲۰۱
توزیع $\bar{X} - \bar{Y}$ ، واریانس دو جامعه معلوم.....	۲۰۱
توزیع $\frac{S_1}{S_2}$	۲۰۲
مجموعه سؤالات کنکوری نمونه‌گیری.....	۲۰۳
پاسخ مجموعه سؤالات کنکوری نمونه‌گیری.....	۲۱۴

فصل نهم: آزمون‌های ناپارامتری

مقدمه.....	۲۲۱
آزمون ناپارامتری نشانه (علامت).....	۲۲۱
آزمون ناپارامتری یو من ویتنی.....	۲۲۲
پیش شرط‌های موردنیاز برای انجام آزمون من- ویتنی.....	۲۲۲
آزمون ناپارامتری کروسکال والیس.....	۲۲۳
آزمون ناپارامتری همبستگی اسپیرمن.....	۲۲۵
آزمون دقیق فیشر.....	۲۲۵
آزمون فریدمن.....	۲۲۶
آزمون کلموگروف- اسمیرانف.....	۲۲۷
آزمون مک نمار.....	۲۲۷
آزمون میانه.....	۲۲۷
آزمون تک نمونه‌ای دورها.....	۲۲۸
آزمون لون Levene.....	۲۲۸
مجموعه سؤالات کنکوری آزمون‌های ناپارامتری.....	۲۳۰
پاسخ مجموعه سؤالات کنکوری آزمون‌های ناپارامتری.....	۲۳۴

فصل دهم: رگرسیون

مقدمه.....	۲۳۷
تعریف.....	۲۳۷
رگرسیون به عنوان متوسط متغیری وابسته به متغیر دیگر.....	۲۳۷
برازش خطی.....	۲۳۷
رگرسیون خطی در حالت نرمال بودن ϵ_i ها.....	۲۳۹
آنالیز واریانس رگرسیون.....	۲۴۰
رگرسیون لجستیک (Logistic Regression).....	۲۴۰
مجموعه سؤالات تألیفی رگرسیون.....	۲۴۲
پاسخ مجموعه سؤالات تألیفی رگرسیون.....	۲۴۵

فصل یازدهم: آمار چند متغیری

مقدمه.....	۲۷۹
تجزیه و تحلیل کواریانس.....	۲۷۹
موارد استفاده از تجزیه کواریانس.....	۲۸۰
مفروضات تجزیه کواریانس.....	۲۸۰
تجزیه و تحلیل کواریانس در طرح بلوک‌های کامل تصادفی.....	۲۸۰
تجزیه و تحلیل کواریانس در طرح مربع لاتین.....	۲۸۱
نحوه محاسبه مقادیر ویژه و بردارهای ویژه.....	۲۸۲
استنباط آماری در مورد میانگین‌ها.....	۲۸۲
فواصل اطمینان هم‌زمان.....	۲۸۴
تجزیه واریانس طرح بلوک‌های کاملاً تصادفی.....	۲۸۴
مؤلفه‌های اصلی.....	۲۸۴
تعداد مؤلفه‌های اصلی.....	۲۸۵
مؤلفه‌های اصلی برای ماتریس کواریانس با ساختارهای ویژه.....	۲۸۵
آزمون ساختار همبستگی‌های مساوی.....	۲۸۸
تحلیل عاملی.....	۲۸۸
مدل آماری تجزیه عاملی.....	۲۸۹
مراحل تجزیه به عامل‌ها.....	۲۹۰
مجموعه سؤالات کنکوری آمار چند متغیری.....	۲۹۱
پاسخ مجموعه سؤالات کنکوری آمار چند متغیری.....	۳۰۲

فصل دوازدهم: شاخص‌های بهداشتی

۳۰۷	مقدمه.....
۳۰۷	شاخص‌های بهداشتی.....
۳۰۸	ابزارهای اندازه‌گیری در اپیدمیولوژی.....
۳۰۸	میزان.....
۳۰۸	انواع مختلف میزان‌ها.....
۳۱۱	تناسب.....
۳۱۱	میزان ابتلا.....
۳۱۱	بروز.....
۳۱۲	شیوع.....
۳۱۲	سایر نکات کنکوری شاخص‌های بهداشتی.....
۳۱۵	خطاها در اپیدمیولوژی.....
۳۱۸	مطالعات تجربی یا مداخله‌ای.....
۳۲۳	مجموعه سؤالات تألیفی شاخص‌های بهداشتی.....
۳۲۹	پاسخ مجموعه سؤالات تألیفی شاخص‌های بهداشتی.....
۳۳۱	پیوست سؤالات ارشد و دکتری ۱۳۹۷.....
۳۳۶	روش‌های آماری (شامل طرح آزمایش‌ها- رگرسیون- نمونه‌گیری- آمار ناپaramتری).....
۳۵۸	فهرست منابع.....

آمار توصیفی



مقدمه

موضوع آمار توصیفی^۱ تنظیم و طبقه‌بندی داده‌ها، نمایش ترسیمی و محاسبه مقادیری از قبیل نما، میانگین، میانه و... می‌باشد که حاکی از مشخصات یکایک اعضای جامعه مورد بحث است. در آمار توصیفی اطلاعات حاصل از یک گروه، همان گروه را توصیف می‌کند و اطلاعات به دست آمده به دسته‌جات مشابه تعمیم داده نمی‌شود. به‌طور کلی از سه روش در آمار توصیفی برای خلاصه‌سازی داده‌ها استفاده می‌شود:

استفاده از جداول

استفاده از نمودار

محاسبه مقادیری خاص که نشان‌دهنده خصوصیات مهمی از داده‌ها باشند.

جامعه آماری

از جامعه تعاریف متعددی به عمل آمده است که به‌طور عمده به این شرح‌اند:

گروهی از افراد که یک یا چند صفت مشترک دارند و این صفت یا صفات، مورد توجه محقق می‌باشند. جامعه ممکن است همه افراد، یک نوع خاص و یا عده محدودتری از همان گروه را شامل شود.

تمامی گروه و افرادی که جهت تحقیق، مناسب تشخیص داده شده‌اند. با این تعریف، جامعه، پسوند «آماري» گرفته است. عده‌ای برای «جامعه‌ی آماری»، تعریف خاص‌تری به این شرح ارائه داده‌اند:

جامعه‌ای که اعضای آن، کمیت‌های مربوط به اندازه‌ی یک صفت در افراد یک جامعه است.

جامعه در تحقیق: مجموعه اجزائی که دارای حداقل یک صفت یا ویژگی مشترک باشند.

جامعه را می‌توان بر اساس معیارهای مختلفی دسته‌بندی نمود که ذیلأً به برخی از آن‌ها اشاره می‌گردد:

بر اساس موضوع که در آن جامعه نوع خود را مشخص می‌کند، مانند جامعه انسانی، جامعه گیاهی و...

بر اساس تعداد اعضای تشکیل‌دهنده، که یا محدود است (می‌توانیم تعداد اعضا یا عناصر تشکیل‌دهنده را مشخص و معین کنیم) و یا نامحدود است (تعداد اعضا یا عناصر تشکیل‌دهنده آن مشخص نیست، مثل درختان یک جنگل).

در تحقیقات، به‌طور معمول سعی می‌شود که جامعه‌ی محدود را مورد مطالعه قرار دهند.

طبقه‌بندی جامعه بر اساس تعمیم که دارای سه نوع است:

الف) جامعه موردنظر: جامعه‌ای است متشکل از همه اعضای واقعی یا فرضی که علاقه‌مند هستیم یافته‌های پژوهش را به آن‌ها تعمیم دهیم. که به آن جامعه‌ی هدف نیز می‌گویند.

^۱- (Descriptive statistics)



ب) جامعه مورد مطالعه: جامعه‌ای است که با استفاده از یک قید از جامعه موردنظر به دست می‌آید و در عمل زیر پوشش بررسی قرار می‌گیرد. این نوع، برای مواردی است که نمی‌توانیم نتایج را به همه افراد جامعه تعمیم دهیم. به این جامعه، جامعه‌ی مورد بررسی نیز می‌گویند.

ج) مطالعه را می‌توان با همه‌ی اعضاء و یا بخشی از اعضای جامعه موردنظر یا جامعه مورد مطالعه به عمل آورد. بخشی از اعضاء را یا به طریق در دسترس و یا به روش نمونه‌گیری علمی می‌توان انتخاب کرد. در طریق اول، جامعه در دسترس و در طریق دوم جامعه نمونه خوانده می‌شود.

جامعه در دسترس: به جامعه‌ای گفته می‌شود که محقق، افراد آن را فقط به خاطر سهولت دسترسی، بدون رعایت اصول نمونه‌گیری علمی، برای مطالعه انتخاب می‌کند.

جامعه نمونه: چون شرح این نوع جامعه مفصل می‌باشد و کاربرد زیاد دارد، آن را تحت عنوان «نمونه» در ادامه مورد شناخت قرار می‌دهیم.

نمونه

ممکن است جامعه‌ی آماری از نظر تعداد افراد یا مواردی که باید مشاهده شود، بزرگ یا کوچک باشد. برای صرفه‌جویی در نیروی انسانی، هزینه و وقت و رعایت سایر ملاحظات اجرایی، به جای مطالعه در مورد تمام افراد جامعه می‌توان نمونه‌ای از افراد جامعه را انتخاب و مورد تحقیق قرار داد. نمونه، معمولاً گروهی از افراد جامعه است که معرف آن جامعه بوده و کم و بیش ویژگی‌های افراد جامعه را داراست. از راه مشاهده و تجزیه و تحلیل اطلاعات مربوط به نمونه، می‌توان بر اساس اصول و قواعد معینی، مشخصات جامعه را استنباط کرد. اندازه‌های به دست آمده درباره‌ی صفات و متغیرهای نمونه را شاخص آماری و اندازه‌هایی را که از روی شاخص آماری نمونه، درباره‌ی جامعه استنباط یا برآورد می‌شود پارامتر می‌نامند.

تعریف نمونه

زیر مجموعه‌ای است از جامعه، که اعضای آن را بخشی از افراد جامعه‌ی اصلی تشکیل می‌دهند. یا در مفهوم وسیع، نمونه عبارت است از «مجموعه‌ای از نشانه‌ها که از یک قسمت یا یک گروه از جامعه‌ی بزرگ‌تر انتخاب شده است، به‌طوری‌که صفات اعضای این مجموعه، معرف کیفیات و ویژگی‌های جامعه‌ی بزرگ‌تر باشد». نمونه، اطلاعاتی را ارائه می‌کند که بر اساس آن می‌توان در مورد خصوصیات جامعه قضاوت کرد؛ به عبارت دیگر، از طریق مطالعه‌ی نمونه و پیدا کردن خصوصیات و روابط پدیده‌ها و تعمیم آن، می‌توان خصوصیات و روابط پدیده‌ها را در جامعه پیدا نمود.

برای اینکه بتوان نتایج به دست آمده از نمونه را به جامعه تعمیم داد، باید دو شرط زیر جمع باشد:

الف) اندازه و حجم نمونه بر اساس منطق و فرمول‌های آماری به دست آید.

ب) روش انتخاب افراد نمونه از بین افراد جامعه، با رعایت موازین علمی باشد.

متغیر

متغیر یک مفهوم است که بیش از دو یا چند ارزش یا عدد به آن اختصاص داده می‌شود. به عبارت دیگر متغیر به ویژگی‌هایی اطلاق می‌شود که می‌توان آنها را مشاهده یا اندازه‌گیری کرد. و دو یا چند ارزش یا عدد را جایگزین آنها نمود. عدد یا ارزش نسبت داده به متغیر، نشان‌دهنده تغییر از یک فرد به فرد دیگر یا از یک حالت به حالت دیگر است. میز یک مفهوم است نه متغیر، اما وزن میز یک متغیر است.

مفهوم میز به تنهایی بر وجود ارزش‌های چندگانه دلالت نمی‌کند و برای مثال، به همین دلیل این مفهوم مشخص نمی‌سازد که چه ویژگی یا ویژگی‌هایی از آن موردنظر است. و برای مثال، کدامیک از ویژگی‌های رنگ، وزن یا ارتفاع میز مورد مشاهده یا اندازه‌گیری قرار گرفته است. کرلینجر (۱۹۸۶) معتقد است متغیر یک نماد است که می‌توان عدد یا ارزش را جایگزین آن کرد. به‌عنوان مثال، X یک متغیر است، و در حقیقت نمادی است که می‌توان عدد یا ارزش را جانشین آن قرار داد.

گاهی اوقات ویژگی‌هایی که در یک پژوهش اندازه‌گیری می‌شوند، ممکن است در پژوهش دیگر ثابت نگه‌داشته شوند. متغیر در مقابل ثابت قرار دارد. ثابت به ویژگی‌هایی اطلاق می‌شود که دارای ارزش مساوی و یکسان است و میزان آن در همه افراد یا اشیاء یا حوادث به یک اندازه است. به‌عنوان مثال، اگر در پژوهشی دانش‌آموزان کلاس چهارم به‌عنوان آزمودنی به کار روند، «کلاس» ثابت است، یا اگر در یک پژوهش دانش‌آموزان ده ساله مشارکت داشته باشند، «سن» ثابت است.

انواع متغیرها

الف) طبقه‌بندی متغیرها برحسب ماهیت متغیر

۱- متغیر کمی و کیفی

متغیر یک مفهوم است که می‌تواند مشاهده یا اندازه‌گیری شود. این اندازه‌گیری ممکن است به‌صورت کیفی یا کمی انجام شود. متغیر کمی به متغیری اطلاق می‌شود که از نظر کمیت تغییر می‌کند و اختلاف مقادیر آن را می‌توان با استفاده از عدد ثبت کرد و آنها را می‌توان با هم جمع کرد. متغیرهای کمی، متغیرهایی هستند که انسان توانسته است برای آنها واحد و مبدأ اندازه‌گیری معین کند، مانند: قد، وزن، سن.

متغیر کیفی، به متغیری اطلاق می‌شود که اختلاف و تغییرات بین میزان‌های مختلف آن کیفی است و برای ثبت آن ممکن است از روش‌های دیگری غیر از به کار بردن عدد استفاده شود. به عبارت دیگر پژوهشگر، توانایی اندازه‌گیری متغیر کیفی را ندارد و ویژگی‌های آن را نمی‌تواند به‌وسیله ارقام ریاضی نمایش دهد. برای ثبت مشاهدات یا اندازه‌گیری‌هایی که از این متغیر به عمل می‌آید، از حروف الفبا یا کد استفاده می‌شود. این‌گونه متغیرها را نمی‌توان جمع و تفریق کرد و برای آنها مبدأ اندازه‌گیری نیز وجود ندارد. رنگ مو، رنگ چشم، و جنس متغیرهای کیفی هستند.

۲- متغیر گسسته و پیوسته

کار پژوهشگر، جمع‌آوری اطلاعات و تجزیه و تحلیل داده‌های جمع‌آوری شده از متغیرهاست. همان‌طور که قبلاً ذکر شد، متغیرها به‌وسیله عدد یا ارزش، مشخص می‌شوند. ماهیت اعداد و ارزش‌ها بستگی به این دارد که متغیر موردنظر پیوسته است یا گسسته.

متغیر گسسته می‌تواند اعداد یا ارزش‌هایی را که مشخص‌کننده یک وجه مشخص و معین از یک مقیاس هستند، به خود اختصاص دهد. به‌عنوان مثال، جنس یک متغیر گسسته است: یک شخص یا زن است یا مرد. اختصاص هر نوع ارزش دیگری بین این دو نوع ارزش امکان‌پذیر نیست. تعداد بازیکنان یک تیم فوتبال نیز یک متغیر گسسته است، زیرا فقط امکان داشتن ۱، ۲، ۳، ۴، ۵، ۶، ۷، ۸، ۹، ۱۰، ۱۱ بازیکن وجود دارد و نه ۷/۵ نفر بازیکن.

متغیر پیوسته، متغیری است که بین دو واحد آن هر نقطه یا ارزشی را می‌توان انتخاب کرد. در این متغیر درجات مختلف اندازه‌گیری وجود دارد و دقت وسیله اندازه‌گیری، تعداد این درجات را تعیین می‌کند. به‌عنوان مثال، وزن یک



متغیر پیوسته است و می‌تواند بین صفر تا بی‌نهایت باشد. قد، زمان، طول یا ارتفاع پرش، درصد چاقی بدن، و سطح هموگلوبین خون متغیرهای پیوسته هستند. ناگفته نماند که در عمل تشخیص بین متغیر پیوسته و گسسته به صورت نظری امکان‌پذیر نیست. دلیل این امر فقدان وسایل اندازه‌گیری دقیق و مناسب است. در خیلی از متغیرهای پیوسته ما ناگزیریم اعداد را به صورت کلی برای اندازه‌گیری به کار ببریم. بهره‌ی هوشی از نظر تئوری یک متغیر پیوسته است. اما در عمل، آزمودنی که برای اندازه‌گیری هوش به کار برده می‌شود، به گونه‌ای است که نمره‌ها را به صورت کلی یا نمره‌های گسسته نشان می‌دهد.

۳- متغیرهای دوارزشی و چندارزشی

متغیرها بر اساس تعداد ارزش‌ها یا عددهایی که به آنها اختصاص داده می‌شود، به دو دسته تقسیم می‌شوند: (الف) دوارزشی، (ب) چند ارزشی.

متغیر دوارزشی به متغیری اطلاق می‌شود که به آن فقط دو ارزش یا دو عدد نسبت داده می‌شود، مانند جنس که دارای دو ارزش زن و مرد است و می‌توان برای ثبت آنها از اعداد صفر و یک یا اعداد دیگری استفاده کرد. کرلینجر (۱۹۸۶) متغیر دوارزشی را دوبخشی نامیده است و چنین می‌نویسد: «برخی از این متغیرها، دو بخشی واقعی هستند مانند زن و مرد، مرگ و حیات و شهری و روستایی که حضور یا عدم حضور ویژگی موجب تقسیم‌بندی آنها می‌شود و برخی از آنها ممکن است دوبخشی ساختگی باشند». متغیر چندارزشی متغیری است که بیش از دو عدد یا دوارزش به آن اختصاص داده می‌شود، مانند سطح تحصیل و هوش که دارای درجات مختلفی هستند و به هر یک از درجات آنها می‌توان عدد یا ارزش معینی را اختصاص داد.

(ب) طبقه‌بندی متغیرها برحسب نقش متغیر در تحقیق

۱- متغیر وابسته

متغیر وابسته، متغیر اصلی مورد توجه محقق است. این متغیر، متغیر پاسخ، برون داد یا ملاک است و عبارت است از وجهی از رفتار یک ارگانیسم که تحریک شده است. متغیر وابسته، مشاهده یا اندازه‌گیری می‌شود تا تأثیر متغیر مستقل بر آن معلوم و مشخص شود. هدف محقق آن است که تغییرپذیری متغیر وابسته را تشریح و پیش‌بینی کند. این متغیر از طریق متغیر مستقل پیش‌بینی می‌شود. از طریق تجزیه و تحلیل متغیر وابسته امکان یافتن پاسخ‌ها یا راه‌حلی برای مسئله ایجاد می‌شود. محقق قصد دارد این متغیرها را و همین‌طور سایر متغیرهای تأثیرگذار بر آن را به صورت کمی و سنجش‌پذیر درآورد.

به عنوان مثال، مدیری را در نظر بگیرید که از میزان فروش محصول جدید که پس از آزمایش بازار (بازارسنجی) عرضه شده ولی از آن جهت که آن‌طور انتظار داشته فروش آن بالا نیست، نگران است. متغیر وابسته در اینجا میزان فروش است. چون میزان فروش می‌تواند متفاوت باشد (می‌تواند کم، متوسط، و زیاد باشد) از این‌رو یک متغیر است، چون میزان فروش عامل اصلی مورد توجه مدیر است، پس متغیر وابسته است.

به عنوان مثالی دیگر، پژوهشگری می‌خواهد ورزش کارکنان نزد مدیران عامل شرکت‌های ایرانی برتر در سال ۱۳۸۸ تا چه اندازه مقبولیت دارد. در اینجا نگرش نسبت به ورزش کارکنان یک متغیر وابسته است. چنین نگرشی می‌تواند دامنه‌ای از کاملاً مخالفم، مخالفم، بی‌تفاوتم، موافقم و کاملاً موافقم در برگیرد.



و در مثالی دیگر، مدیر منابع انسانی شرکتی از این جهت که کارکنان نسبت به سازمان وفادار نیستند، و در حقیقت وفاداری‌های آنها در حال تغییر جهت به سمت سایر مؤسسات است، احساس نگرانی می‌کند. در این مورد، وفاداری سازمانی، متغیر وابسته است.

در اینجا نیز، در سطوح وفاداری سازمانی کارکنان، نوعی پراکندگی وجود دارد. مدیر منابع انسانی ممکن است بخواهد بداند چه عواملی در وفاداری اعضاء سازمانی نقش دارند، طوری که بتواند این میزان نوسان را در وفاداری کنترل کند. فرض کنید پی می‌برد که میزان پرداخت بالا می‌تواند وفاداری اعضاء نسبت به سازمان را حفظ کند و در نتیجه، آنها شغل خود را برای دستیابی به شغل دیگر در سازمانی دیگر، ترک نخواهند کرد. افزایش میزان پرداخت به کارکنان ممکن است، به کنترل میزان تغییرپذیری در وفاداری کمک کند و افراد ممکن است، این سازمان را ترک نکنند. ممکن است، در یک پژوهش بیش از یک متغیر وابسته وجود داشته باشد. برای مثال، همیشه بین کیفیت و حجم ستاده، و تولید با هزینه پایین نوعی تنش وجود دارد. در چنین مواردی، مدیر مایل است به درک و شناخت عواملی که بر همه متغیرهای وابسته موردنظر تأثیر می‌گذارند، و چگونگی تفاوت برخی از این عوامل در مورد برخی از متغیرهای وابسته آگاه شود. این بررسی‌ها، ممکن است مستلزم تجزیه و تحلیل آماری چند متغیره (رگرسیون چندگانه) باشد.

۲- متغیر مستقل

متغیر مستقل، متغیر محرک درونداد است که به وسیله پژوهشگر اندازه‌گیری، دستکاری یا انتخاب می‌شود تا تأثیر یا ارتباط آن با متغیر دیگری معین شود. متغیر مستقل پیش‌فرض متغیر وابسته است. به عبارت دیگر این متغیر، مقدمه و متغیر وابسته، نتیجه است. در تحقیق آزمایشی، متغیر مستقل به وسیله آزمایش‌کننده دستکاری می‌شود تا تأثیر تغییرات آن بر متغیر دیگری که وابسته فرض شده است، مشاهده و اندازه‌گیری شود. به عبارت دیگر، تغییرپذیری یا نوسان در متغیر وابسته به حساب متغیر مستقل گذاشته می‌شود.

به عنوان مثال، فرض کنید که بررسی‌های تحقیقاتی نشان می‌دهد که محصول جدید موفقیت‌آمیز بوده است. در اینجا تولید موفقیت‌آمیز محصول جدید، قیمت بازاری سهام را تحت تأثیر قرار می‌دهد و نوسان در آن را نشان می‌دهد. یعنی، موفقیت بیشتر محصول جدید، قیمت بازاری سهام آن شرکت را بالاتر خواهد برد. به علاوه، موفقیت محصول جدید، متغیر مستقل و قیمت بازاری سهام، متغیر وابسته است. میزان موفقیت محصول جدید نوسان در قیمت بازاری سهام شرکت را تبیین خواهد کرد. این رابطه و عناوین متغیرها در شکل زیر نشان داده شده است.

به عنوان مثال دیگر، فرض کنید تحقیقات فرهنگی نشان می‌دهد که ارزش‌های مدیریتی، فاصله قدرت بین مافوق‌ها و زیردستان را تعیین می‌کند. در اینجا، فاصله قدرت (یعنی روابط بین رئیس و زیردست و کارکنان در مقایسه با رئیس پر قدرت در مقابل در تعامل با رئیس کم قدرت) موضوع مورد توجه است، لذا متغیر وابسته است. ارزش‌های مدیریتی که تغییرپذیری در فاصله قدرت را تبیین می‌کند، متغیر مستقل است. این رابطه در شکل زیر نشان داده شده است.

۳- متغیر تعدیل‌کننده

متغیر تعدیل‌کننده، متغیری است که بر رابطه متغیر مستقل و متغیر وابسته تأثیر اقتضایی دارد. یعنی، حضور متغیر سوم (متغیر تعدیل‌کننده) رابطه مورد انتظار اصلی بین متغیرهای مستقل و وابسته، را تغییر می‌دهد. این متغیر در مثال‌های زیر روشن می‌شود.





فرض کنید پی برده شد که بین تعداد کتاب‌هایی که کودکان ۵ تا ۶ ساله در منزل به آن دسترسی دارند و توانایی‌های خواندن آنها نوعی رابطه وجود دارد. یعنی اگر کودکان ۵ تا ۶ ساله، تعداد زیادی کتاب در اختیارشان قرار گیرد، مهارت‌ها و توانایی خواندن آنها بهبود خواهد یافت.

چون کودکان برای خواندن کتاب‌های بیشتر، فرصت‌های بیشتری دارند (فعالیتی که در آن از جانب والدین کمک شده‌اند) و بنابراین، بهتر می‌خوانند. در نتیجه، اگر کودکان در منزلی که هیچ نوع کتابی وجود ندارد، پرورش داده شوند، فرصتی برای پرورش نحوه خواندن نخواهند داشت، و بنابراین مهارت‌ها و توانایی‌های خواندن در آنها کاملاً شکل نخواهد گرفت. از این رو، می‌توان گفت که نوعی رابطه بین متغیر مستقل (تعداد کتاب‌ها) و متغیر وابسته (توانایی‌های خواندن) وجود دارد که در شکل زیر نشان داده شده است.

اگرچه این رابطه می‌تواند به‌طور کلی برای همه کودکان صدق کند، با این وجود، سطح سواد والدین باعث می‌شود که کودکان قبل از ورود به مدرسه خواندن را بیاموزند، از این رو وجود این پدیده، بر رابطه فوق‌الذکر تأثیر می‌گذارد. در منزلی که والدین بی‌سوادند، هر مقدار کتاب در منزل باشد به توسعه و تکامل مهارت‌ها و توانایی‌های کودکان کمک نخواهد کرد. میزان سواد والدین رابطه بین تعداد کتاب‌ها و توانایی‌های خواندن را تعدیل می‌کند. در موقعیت‌های مختلف، رابطه بین تعداد کتاب‌هایی که کودکان در منزل در اختیار دارند و توانایی‌های کودکان ۵ تا ۶ ساله وابسته و متکی به سواد والدین است. این میزان تأثیرگذاری بر رابطه بین متغیر مستقل و وابسته می‌تواند در شکل زیر نشان داده شود.

به مثابه مورد بالا، هرگاه رابطه بین متغیر مستقل و متغیر وابسته متکی به متغیر دیگری شود، می‌گوییم متغیر سوم، نوعی اثر تعدیل‌کننده روی رابطه متغیر مستقل و وابسته دارد. متغیری که رابطه را تعدیل می‌کند به‌عنوان متغیر تعدیل‌کننده مشهور است.

اجازه دهید مثالی دیگر از متغیر تعدیل‌کننده مرتبط به محیط سازمانی ارائه دهیم. یکی از تئوری‌های جدید مدعی است که تنوع نیروی کار (مشمول بر مبانی اخلاقی، نژادها و ملیت‌های مختلف) در اثربخشی سازمانی نقش ایفا خواهد کرد زیرا هر گروه، مهارت‌ها و تخصص‌های فنی خاصی را با خود به محیط کاری می‌آورد. این هم‌افزایی می‌تواند محقق شود، فقط اگر مدیران بدانند چگونه استعدادهای خاص گروه کاری متنوع را به کار ببندند. در سناریو بالا، اثربخشی سازمانی متغیر وابسته است که به‌وسیله نیروی کاری متنوع به‌طور مثبت تحت تأثیر قرار می‌گیرد، که در واقع متغیر مستقل است. در عین حال، برای استفاده از نیروهای بالقوه، مدیران باید بدانند چگونه استعدادهای گروه‌های مختلف را برای اثربخش ساختن امور، تشویق و هماهنگ سازند. اگر ندانند، هم‌افزایی محقق نخواهد شد. به عبارت دیگر، بهره‌برداری اثربخش از استعدادها، دیدگاه‌های مختلف و توانایی‌های حل مسئله برای افزایش اثربخشی سازمانی متکی و وابسته به مهارت مدیران در عمل کردن به‌عنوان، یک عامل تسریع‌کننده است. این مهارت فنی مدیریتی به‌صورت متغیر تعدیل‌کننده در می‌آید. این روابط در شکل زیر نشان داده شده است.

تفاوت بین متغیر مستقل و متغیر تعدیل‌کننده

ممکن است فرد در تمیز دادن این که چه موقع متغیری مستقل و چه موقع متغیری تعدیل‌گر است، دچار سردرگمی و ابهام شود. برای مثال ممکن است با دو موقعیت زیر مواجه باشیم:

موقعیت اول

یکی از تحقیقات انجام شده نشان می‌دهد که اگر کیفیت برنامه‌های آموزشی در یک سازمان بهتر بوده و نیاز به رشد کارکنان نیز زیاد باشد، افراد تمایل بیشتری به آموختن روش‌های جدید انجام کارها دارند.

موقعیت دوم

یکی دیگر از تحقیقات نشان می‌دهد که تمایل کارکنان به آموختن روش‌های انجام کار، تحت تأثیر کیفیت برنامه‌های آموزشی ارائه شده به‌وسیله سازمان (که برای همه افراد سازمان اجرا و برگزار می‌شود) قرار نمی‌گیرد. فقط کارکنانی که به نظر می‌رسد نیاز به رشد بالایی دارند، تمایل به آموختن روش‌های انجام کار از طریق آموزش‌های تخصصی دارند.

در دو موقعیت بالا، همان سه متغیر را داریم. در مورد اول، برنامه‌های آموزشی و شدت نیاز رشد، متغیرهای مستقلی‌اند که تمایل کارکنان به آموختن (متغیر وابسته) را تحت تأثیر قرار می‌دهند. در مورد دوم، کیفیت برنامه‌های آموزشی متغیر مستقل است در حال که متغیر وابسته همان اولی است ولی شدت نیاز به رشد به‌صورت متغیر تعدیل‌کننده درمی‌آید. به عبارت دیگر، وقتی کیفیت برنامه آموزشی افزایش داده می‌شود، فقط کسانی که نیاز به رشد بالایی دارند (برای مثال نیاز به رشد قوی) تمایل بیشتری به آموختن روش‌های انجام کار از خود نشان می‌دهند. بنابراین، در این حالت رابطه بین متغیرهای مستقل و متغیرهای وابسته متکی و وابسته به فلسفه وجودی یک متغیر تعدیل‌گر است.

موارد فوق‌الذکر، این موضوع را روشن می‌سازد حتی اگرچه ممکن است متغیرهای مورد استفاده مشابه باشند، تصمیم‌گیری در مورد عنوان‌گذاری متغیر به‌صورت مستقل، وابسته و تعدیل‌کننده می‌تواند بستگی به نحوه تأثیرگذاری آنها بر متغیر دیگر باشد. تفاوت‌های بین اثرات متغیرهای مستقل و تعدیل‌کننده در نمودارهای زیر نشان داده شده است.

نمودار «الف» تأثیر متغیرهای مستقل روی متغیر وابسته، زمانی که هیچ‌گونه متغیر تعدیل‌کننده در موقعیت فعالیت نمی‌کند، را نشان می‌دهد. و نمودار «ب» تأثیر متغیر مستقل بر متغیر وابسته وقتی که متغیر تعدیل‌کننده در وضعیت عمل می‌کند را نشان می‌دهد.

۴- متغیر مداخله‌گر

متغیر مداخله‌گر، متغیری است که از زمانی که متغیرهای مستقل شروع به تأثیرگذاری بر متغیر وابسته می‌کنند و تا زمان این تأثیرگذاری، ظاهر می‌شود. بنابراین، نوعی حالت موقت یا بعد زمانی برای متغیر مداخله‌گر وجود دارد. متغیر مداخله‌گر به‌عنوان تابعی از متغیر(های) مستقل که در هر وضعیتی عمل می‌کند، آشکار و ظاهر می‌گردد و به مفهوم‌سازی و تشریح تأثیر متغیر(های) مستقل بر متغیر وابسته کمک می‌کند. مثال زیر این نکته را تشریح می‌کند. در یکی از مثال‌های قبل جایی که، متغیر مستقل تنوع نیروی کاری، متغیر وابسته اثربخشی سازمانی را تحت تأثیر قرار داد، یک متغیر مداخله‌گر به‌عنوان تابعی از تنوع در نیروی کار ظاهر می‌شود هم‌افزایی خلاقانه است. این هم‌افزایی خلاقانه، از نوعی نیروی کار متفاوت، چند نژادی و چند ملیتی (مثلاً تنوع) که با هم در تعامل‌اند و مهارت‌های فنی چندگانه خود را در حل مسئله به درون گروه می‌آورند نشأت می‌گیرد. این امر ما را کمک می‌کند تا درک نماییم چگونه اثربخشی سازمانی می‌تواند از تنوع در نیروی کار ناشی شود. توجه کنید که هم‌افزایی خلاقانه، که همان متغیر مداخله‌گر است در زمان T_2 به‌عنوان تابعی از تنوع نیروی کاری ظاهر می‌شود تا T_1 و اثربخشی سازمانی را در زمان T_3 با خود به وجود می‌آورد. متغیر مداخله‌گر هم‌افزایی خلاقانه به ما کمک می‌کند تا درک کنیم چگونه تنوع نیروی کاری، اثربخشی سازمانی را با خود به وجود می‌آورد. پویایی‌های این رابطه در شکل زیر نشان داده شده است. این شکل روابط بین متغیرهای مستقل، مداخله‌گر و وابسته را نشان می‌دهد.

رابطه بین متغیرها

جالب است ببینیم چگونه ورود متغیر تعدیل کننده (مهارت فنی مدیریتی) در مثال قبلی می‌توانست مدل را تغییر داده یا روابط را تحت تأثیر قرار دهد. مجموعه جدیدی از روابط که می‌تواند با وجود متغیر تعدیل کننده ظاهر شوند. همان طور که می‌بینید مهارت فنی مدیریتی، رابطه بین تنوع نیروی کاری و هم‌افزایی خلاقانه را تعدیل می‌کند. به عبارت دیگر، هم‌افزایی خلاقانه از مهارت‌های چندگانه حل مسئله نیروی کار متنوع تا زمانی که مدیر، شایستگی ایجاد هم‌افزایی از طریق هماهنگی خلاقانه مهارت‌های مختلف نداشته باشد، ظاهر نخواهد شد. اگر مدیر هیچ نوع مهارتی برای انجام این نقش نداشته باشد و آنگاه نحوی بهره‌برداری از مهارت‌های متعدد حل مسئله نیروی کاری متنوع حائز اهمیت نبوده، در نتیجه هم‌افزایی حاصل نمی‌شود. بنابراین، سازمان ممکن است به جای اینکه به صورت اثربخش کار کند، به حالت ایستا باقی بماند.

اکنون به سادگی می‌توان تفاوت‌هایی که بین متغیر مستقل، مداخله‌گر و تعدیل کننده وجود دارد را مشاهده کرد. متغیر مستقل نوسان یا تغییر در متغیر وابسته را تبیین می‌کند؛ و متغیر مداخله‌گر در زمان T_2 به عنوان تابعی از متغیرهای مستقل و متغیرهای وابسته ظاهر می‌شود که به ما کمک می‌کند تا رابطه بین متغیرهای مستقل و وابسته را مفهوم‌سازی کنیم؛ و متغیر تعدیل کننده، اثری عارضی (محتمل الوقوع) بر رابطه بین دو متغیر دارد. در شرایط مختلف، زمانی که متغیر مستقل، نوسان در متغیر وابسته را تشریح و توجیه می‌کند، متغیر مداخله‌گر به نوسانی که قبلاً به وسیله متغیر مستقل تبیین شده چیزی اضافه نمی‌کند، در حالی که متغیر تعدیل کننده و متغیر مستقل در تشریح نوسان یا واریانس، دارای نوعی اثر تعاملی‌اند. یعنی، وقتی متغیر تعدیل کننده حضور نداشته باشد رابطه فرض شده بین دو متغیر پایدار نمی‌ماند. خواه یک متغیر، متغیر مستقل باشد خواه متغیر وابسته یا مداخله‌گر یا متغیر تعدیل کننده، باید نوع آن، با مطالعه دقیق عواملی که در هر وضعیت عمل می‌کنند تعیین شوند. برای مثال، یک متغیر نظیر انگیزه کارکنان می‌تواند بر اساس مدل نظری که قبلاً مطرح شد، یک متغیر وابسته، مستقل، مداخله‌گر یا یک متغیر تعدیل کننده باشد.

جدول توزیع فراوانی دسته‌بندی نشده

این طریقه تنظیم داده‌ها اغلب برای داده‌های جدا استفاده می‌شود فرض کنید n داده داشته باشیم که k تا از آنها ($k \leq n$) متمایز باشد ابتدا مقادیر متمایز داده‌ها را به طور صعودی مرتب کرده و در یک ستون جدول قرار می‌دهیم و مقابل آنها فراوانی‌های مربوطه را می‌نویسیم:

x_i	f_i	r_i	F_i	F'_i
x_1	f_1	r_1	F_1	F'_1
x_2	f_2	r_2	F_2	F'_2
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
x_k	f_k	r_k	F_k	F'_k
	n	1		

f_i فراوانی مطلق مربوط به مقدار X_i تعداد دفعاتی را نشان می‌دهد که X_i در داده‌ها مشاهده شده و r_i فراوانی نسبی مربوط به مقدار X_i برابر است با $\frac{f_i}{n}$ از این رو داریم $\sum_{i=1}^k r_i = 1$ و $\sum_{i=1}^k f_i = n$ F_i فراوانی تجمعی (انباشته) مربوط به X_i برابر است با تعداد داده‌های برابر یا کمتر از X_i بنابراین:

$$F_1 = f_1$$

$$F_2 = f_1 + f_2$$

⋮

$$F_i = f_1 + f_2 + \dots + f_{i-1} + f_i$$

⋮

$$F_k = \sum_{i=1}^k f_i = n$$

F'_i فراوانی تجمعی نسبی مربوط به مقدار X_i برابر است با $\frac{F_i}{n}$ از این رو:

$$F'_1 = r_1$$

$$F'_2 = r_1 + r_2$$

⋮

$$F'_i = r_1 + r_2 + \dots + r_{i-1} + r_i$$

⋮

$$F'_k = 1$$

مثال ۱-۱: تعداد افراد ۲۰ خانوار عبارت‌اند از:

۵،۶،۵،۵،۲،۵،۶،۲،۳،۳،۲،۲،۵،۷،۴،۴،۴،۲،۳

X_i	f_i	r_i	F_i	F'_i
۲	۶	%۳۰	۶	%۳۰
۳	۳	%۱۵	۹	%۴۵
۴	۳	%۱۵	۱۲	%۶۰
۵	۵	%۲۵	۱۷	%۸۵
۶	۲	%۱۰	۱۹	%۹۵
۷	۱	%۵	۲۰	%۱۰۰
	۲۰	۱		

جدول توزیع فراوانی دسته‌بندی شده

اطلاعات به دست آمده از یک تحقیق غالباً توده‌ای از اطلاعات خام، بر معنی و بدون نظم هستند که هر نوع نتیجه‌گیری و تفسیر آنها غیرممکن است. بنابراین برای هر نوع تجزیه و تحلیل اطلاعات لازم است داده‌ها (بخصوص



داده‌هایی که در سطح مقیاس اندازه‌گیری فاصله‌ای و نسبی به دست آمده‌اند) براساس یک نظم منطقی طبقه‌بندی^۱ شوند تا به صورت معنی‌دار و قابل تفسیر در آید. طبقه‌بندی داده‌ها مستلزم محاسبه مرحله به مرحله دامنه تغییرات، تعداد طبقات، فاصله طبقات، انواع فراوانی‌ها با استفاده از فرمول‌های مشخص است. طبقه‌بندی داده‌ها تمام اطلاعات در یک جدول به نام جدول توزیع فراوانی^۲ گردآوری می‌شود و این جدول باید اساسی برای محاسبه شاخص‌های مرکزی^۳، شاخص‌های پراکندگی^۴ و مقایسه گروهی از داده‌ها با گروه‌های دیگر جهت استنباط آماری است.

قواعد کلی برای تشکیل جدول توزیع فراوانی طبقه‌بندی شده

اگر تفاضل بزرگ‌ترین مشاهده از کوچک‌ترین مشاهده در بین داده‌های آماری بزرگ‌تر از ۲۰ باشد برای طبقه‌بندی و تنظیم داده‌ها از جدول توزیع فراوانی طبقه‌بندی شده استفاده می‌شود. برای تنظیم این جدول به شرح زیر اقدام می‌شود:
۱- تعیین دامنه تغییرات با استفاده از فرمول زیر:

$$R = X_{\max} - X_{\min}$$

۲- تعیین تعداد طبقات با استفاده از یکی از روابط دلخواه زیر:

$$K = \sqrt{n} \quad K = 1 + 3.3 \log n$$

البته انتخاب تعداد طبقات اختیاری است ولی معمولاً تعداد طبقات را بین ۲۰-۵ انتخاب می‌نمایند. طبقات بیشتر از ۲۰ محاسبات را طولانی می‌نماید و انتخاب طبقات کمتر از ۵ باعث از دست رفتن اطلاعات می‌شود. اگر در محاسبه طبقات، عدد حاصل اعشاری آمد همواره به طرف بالا گرد می‌نمایند حتی اگر قسمت اعشاری کمتر از ۰/۵ باشد.
۳- تعیین طول دسته‌ها با استفاده از فرمول زیر:

$$C = \frac{R}{K}$$

R: دامنه تغییرات

K: تعداد طبقات

۴- مشخص کردن حدود رده‌ها از طریق روابط زیر:

طول رده + حد پایین رده اول = حد پایین رده دوم

طول رده + حد بالای رده اول = حد بالای رده دوم

۵- نوشتن طبقات:

معمولاً نوشتن طبقات را از کوچک‌ترین عددی که اندازه یا فاصله طبقات مضربی از آن است آغاز می‌نمایند همچنین می‌توان نوشتن تعداد طبقات را از کوچک‌ترین عدد شروع کرد.

۶- تعیین تعداد فراوانی هر طبقه در ستون دوم جدول

۷- تشکیل فراوانی‌های نسبی در ستون سوم جدول

۸- وارد کردن نماینده وسط طبقات در ستون چهارم جدول

^۱- Classification

^۲- Frequency Table

^۳- Central Index

^۴- Dispersion Index



۹- تشکیل فراوانی‌های تراکمی یا طاقوار در ستون پنجم جدول

۱۰- تشکیل فراوانی‌های تراکمی یا طاقوار درصدی در ستون ششم جدول

۱۱- تشکیل ستون حدود واقعی طبقات (کرانه طبقات) در ستون هفتم جدول

مثال ۱-۲: محقق سطح هوشی یک گروه ۱۰۵ نفری از دانش آموزان را با استفاده از آزمون هوشی و کسلر برای کودکان اندازه‌گیری کرده است. نمرات به دست آمده دارای پراکندگی بسیاری هستند و به دست آوردن اطلاعات لازم از قبیل «درصد دانش‌آموزانی که در سطح هوشی معینی قرار دارند» غیرممکن است. حتی پیدا کردن بالاترین و پایین‌ترین سطح هوشی به سختی امکان‌پذیر است. بنابراین برای معنی‌دار و قابل تفسیر شدن داده‌ها لازم است نمرات هوشی طبقه‌بندی شوند. این طبقه‌بندی در مراحل زیر انجام می‌گیرد:

۹۰	۱۰۰	۱۰۰	۸۰	۱۱۲	۱۳۰	۷۰	۷۰	۸۶	۸۹	۹۰	۹۵	۱۳۰	۶۵	۶۵	۷۰	۱۰۰	۱۱۰
۱۰۰	۹۰	۸۰	۹۰	۸۵	۱۱۲	۱۱۵	۸۵	۷۵	۹۰	۹۵	۱۰۵	۹۵	۱۲۰	۱۰۵	۹۵	۷۵	۶۵
۱۰۰	۱۲۰	۱۱۰	۱۱۵	۱۰۵	۸۵	۸۰	۶۰	۸۰	۷۰	۱۰۰	۱۰۰	۱۱۰	۱۲۰	۱۱۱	۱۱۵	۹۰	۱۰۰
۹۵	۸۵	۸۰	۱۱۰	۱۱۰	۹۵	۱۰۵	۱۰۰	۸۰	۹۹	۱۱۵	۱۲۵	۹۰	۱۰۵	۱۲۵	۱۰۵		۱۰۰
۹۵	۸۵	۷۵	۷۰	۷۵	۱۱۰	۱۰۰	۱۰۰	۹۰	۹۵	۱۰۵	۱۱۰	۱۰۵	۱۰۵	۸۵	۷۰		۶۰
۶۵	۱۰۰	۱۱۰	۹۰	۸۵	۹۰	۸۰	۱۰۰	۱۱۰	۱۱۵	۱۰۵	۱۳۰	۱۲۵	۱۰۰	۹۰	۹۵		۱۰۵

محاسبه دامنه تغییرات

اگر متغیر مورد اندازه‌گیری (مانند هوش) را با حذف "X" نشان دهیم دامنه تغییرات با استفاده از فرمول مثال یک (تفاضل بزرگ‌ترین عدد از کوچک‌ترین عدد به اضافه یک) به دست می‌آید.

دامنه تغییرات سطح هوشی دانش آموزان:

$$R = 130 - 60 + 1 = 71$$

محاسبه تعداد طبقات

روش تجربی

در این روش تعیین تعداد طبقات در اختیار محقق است ولی معمولاً آن را بین ۱۰ تا ۲۰ طبقه انتخاب می‌کنند. چون طبقات کمتر از ۱۰ باعث بزرگ‌تر شدن اندازه طبقات و از دست رفتن اطلاعات می‌شود و طبقات بالاتر از ۲۰ باعث طولانی شدن تهیه و تنظیم جدول می‌شود. در مثال فوق محقق تعداد طبقات به روش تجربی "۱۵۰" طبقه در نظر گرفت.

روش فرمولی

در روش فرمولی تعداد طبقات از طریق فرمولی زیر که به قانون استرژنبر معروف است به دست می‌آید.

گاریتم بر مبنای ۱۰: $\text{Log: } n = \text{تعداد اعداد} = \text{تعداد طبقات } K$

$$k = 1 + \sqrt[3]{\log n}$$

محاسبه فاصله طبقات

محاسبه طبقات از تقسیم دامنه تغییرات بر تعداد طبقات از طریق فرمول زیر به دست می‌آید.

$$C = \frac{R}{K}$$



$$c = \frac{71}{15} = 4.73 \approx 5$$

نوشتن طبقات

معمولاً نوشتن طبقات را از پایین و با عددی شروع می‌کنند که فاصله طبقات مضربی از آن باشد. در مثال فوق کوچک‌ترین عدد ۶۰ است و فاصله طبقات "۵" بنابراین اولیه طبقه با ۶۰ شروع می‌شود (۶۰ مضربی از ۵ است) و به ۶۴ ختم می‌شود (بین "۶۰" تا "۶۴" پنج عدد ۶۰-۶۱-۶۲-۶۳-۶۴ قرار دارد). پس از نوشتن اولین طبقه سایر طبقات را به همان ترتیب می‌نویسند تا به آخرین طبقه برسند.

حدود واقعی طبقات	x_c	P_d	p	P_f	cf	f	طبقات
۱۲۹,۵-۱۳۴,۵	۱۳۲	۱۰۰	۳	۰,۰۳	۱۰۵	۳	۱۳۰-۱۳۴
۱۲۴,۵-۱۲۹,۵	۱۲۷	۹۷,۱	۳	۰,۰۳	۱۰۲	۳	۱۲۵-۱۲۹
۱۱۹,۵-۱۲۴,۵	۱۲۲	۹۴,۳	۳	۰,۰۳	۹۹	۳	۱۲۰-۱۲۴
۱۱۴,۵-۱۱۹,۵	۱۱۷	۹۱,۴	۵	۰,۰۵	۹۶	۵	۱۱۹-۱۱۵
۱۰۹,۵-۱۱۴,۵	۱۱۲	۸۶,۶	۱۱	۰,۱۱	۹۱	۱۲	۱۱۰-۱۱۴
۱۰۴,۵-۱۰۹,۵	۱۰۷	۷۵,۲	۱۰	۰,۱۰	۷۹	۱۱	۱۰۵-۱۰۹
۹۹,۵-۱۰۴,۵	۱۰۲	۶۵	۱۴	۰,۱۴	۶۸	۱۵	۱۰۰-۱۰۴
۹۴,۵-۹۹,۵	۹۷	۵۰,۵	۹	۰,۰۹	۵۳	۱۰	۹۵-۹۹
۸۹,۵-۹۴,۵	۹۲	۴۱	۱۰	۰,۱۰	۴۳	۱۱	۹۰-۹۴
۸۴,۵-۸۹,۵	۸۷	۳۰,۵	۸	۰,۰۸	۳۲	۹	۸۵-۸۹
۷۹,۵-۸۴,۵	۸۲	۲۲	۷	۰,۰۷	۲۳	۷	۸۰-۸۴
۷۴,۵-۷۹,۵	۷۷	۱۵,۲	۴	۰,۰۴	۱۶	۴	۷۹-۷۵
۶۹,۵-۷۴,۵	۷۲	۹,۵	۶	۰,۰۶	۱۲	۶	۷۰-۷۴
۶۴,۵-۶۹,۵	۶۷	۵,۷	۴	۰,۰۴	۶	۴	۶۵-۶۹
۵۹,۵-۶۴,۵	۶۲	۲	۲	۰,۰۲	۲	۲	۶۰-۶۴
			۱۰۰	۱	$N=105$		