بسم الله الرحمن الرحیم

موسسه آموزش عالی آزاد
ماهان
www.mahan.ac.ir

آمار و اقتصاد سنجی

دکتری

مجموعه

اقتصاد

مؤلفان: ابوالقاسم مرآت

قدرت ا... اماموردی

مرآت، ابوالقاسم
آمار و اقتصاد سنجی رشته اقتصاد / ابوالقاسم مرآت، قدرت ا.. امام‌وردی
مشاوران صعود ماهان: ۱۴۰۱
۲۲۴ص: جدول، نمودار (آمادگی آزمون دکتری اقتصاد)
ISBN: 978-600-458-645-0
فهرستنویسی بر اساس اطلاعات فیپا.
فارسی ـ چاپ اول
۱- آمار واقتصاد سنجی ۲- آزمونها و تمرینها (عالی) ۳- آزمون دوره‌های تحصیلات تکمیلی
٤- دانشگاهها و مدارس عالی ـ ایران ـ آزمونها
ابوالقاسم مرآت
ج ـ عنوان
کتابخانه ملی ایران ٤٤٢٠٠٨٩

❒ نام کتاب: .... آمار و اقتصاد سنجی
❒ مولفان: .... ابوالقاسم مرآت، قدرت ا... امام وردی
❒ مدیران مسئول: .... هادی و مجید سیاری
❒ برنامه‌ریزی محتوا: .... سمیه بیگی
❒ ناشر: .... مشاوران صعود ماهان
❒ نوبت و تاریخ چاپ: .... چاپ اول /۱۴۰۱
❒ تیراژ: .... ۱۰۰۰ نسخه
❒ قیمت: .... ۲/۵۹۰/۰۰۰ ریال
❒ شابک: ISBN۹۷۸-۶۰۰-۴۵۸-۶۴۵-۰

انتشارات مشاوران صعود ماهان: تهران ـ خیابان ولیعصر، بالاتر از تقاطع ولیعصر مطهری، پلاک ۲۰۵۰
تلفن: ۸۸۱۰۰۱۱۳ و ۸۸۴۰۱۳۱۳

بنام خدا

ایمان دارم که هر تغییر و تحول بزرگی در مسیر زندگی بدون تحول معرفت و نگرش میسر نخواهد بود. پس بیایید با اندیشه توکل، تفکر، تلاش و تحمل در توسعه دنیای فکریمان برای نیل به آرامش و آسایش توأمان اولین گام را برداریم. چون همگی یقین داریم دانایی، توانایی می‌آورد.

شاد باشید و دلی را شاد کنید

برادران سیاری

مقدمه مولف

با توجه به اهمیت و نقش آمار در برنامه‌ریزیهای توسعه اقتصادی، اجتماعی و فرهنگی و نیز کاربرد آن در رشته‌های مختلف علوم، آشنائی دانش‌پژوهان با اصول و مفاهیم و تکنیکهای پایه‌ای تحلیل آماری در اکثر دوره‌های تحصیلی ضروری به نظر می‌رسد.
مجموعه حاضر بر اساس نیاز داوطلبان شرکت در آزمون های دوره دکترا در رشته‌های اقتصاد مدیریت در گرایشهای مختلف تهیه شده است و کوشش گردیده مطالب به اختصار مطرح گردد.
کلیه مطالب مورد نیاز در سه بخش عمده ارائه گردیده است:
بخش اول: آمار توصیفی می‌باشد که موضوع آن بیان نتایج عینی تحقیق است.
بخش دوم: نظریه احتمال می‌باشد که موضوع آن، مفهوم احتمال کمیتهای تصادفی و قوانین توزیع آنها و مشخص‌کننده‌های عددی کمیتهای تصادفی می‌باشد.
بخش سوم: آمار استنتاجی است که موضوع آن بیان تفسیر و تعیین میزان اعتبار یافته‌های تحقیق می‌باشد. از آنجا که آشنائی با مفاهیم بخش اول و بخش دوم که پیش نیاز بخش سوم می‌باشد، ضروری بنظر می‌رسد، هر چند امکان دارد اکثر دانش‌پژوهان دردوره‌های تحصیلی، این مفاهیم را فرا گرفته باشند. ولی برای آن دسته از عزیزان که آنها را فراموش کرده‌اند، از جهت یادآوری، به اختصار به آنها اشاره شده است. بخش اول در سه قسمت مجزا تحت عنوان، مفاهیم اساسی آمار، مشخص‌کننده‌های عددی و توزیعهای دو بعدی ارائه می‌گردد.
بخش دوم، قضایای مربوط به احتمال، کمیتهای تصادفی و قوانین توزیع آنها، قضایای حدی، مورد بحث قرار می‌گیرد.
در بخش سوم، نظریه نمونه، تخمین زنها، آزمون فرضیه‌های آماری و تحلیلی همبستگی و رگرسیون مطرح شده است.
اگر چه مجموعه حاضر، سهم زیادی از نیازهای علاقمندان به مفاهیم و تکنیکهای آماری را برآورده می‌کند، معذالک کاستی‌هایی در آن نیز وجود دارد که بعلت کمبود وقت، رفع آنها به زمان مناسب تری موکول گردید. از استفاده کنندگان از این کتاب تقاضا می‌شود چنانچه کمبودهایی در مطالب آورده شده احساس می‌کنند. نگارنده را در رفع آنها در چاپ بعدی یاری رسانند.
در خاتمه از مدیریت محترم مؤسسه آموزش عالی ماهان کمال تشکر حاصل است که امکانات چاپ این مجموعه را فراهم نموده‌اند، همچنین از مدیریت انتشارات موسسه آموزش عالی ماهان جناب آقای سید علی ونکی در مساعدت چاپ این کتاب سپاسگزار می‌باشم.

فهرست مطالب

بخش اول: آمار توصیفی

بخش دوم : نظریه احتمال

فهرست مطالب

فهرست مطالب

بخش اول

آمار توصیفی

آمار توصیفی

۱- مفاهیم اساسی آمار: علم آمار نیز مانند سایر علوم بر پایه یک سری از مفاهیم بنا شده است که به مفاهیم اولیه معروفند و بدون تعریف، پذیرفته شده و بصورت زیر توصیف می‌شوند:

- **جامعه آماری:** به کلیه عناصر، افراد و یا اشیاء که حداقل در یک خاصیت مشترک باشند، جامعه آماری گویند.

- **صفت مشخصه:** به خاصیت یا خواصی که در همه عناصر، افراد یا اشیاء مشترک است، صفت مشخصه گفته می‌شود.

- **صفت متغیر:** خاصیت یا خواصی که از یک عنصر به عنصر دیگر، از یک فرد به فرد دیگر جامعه تغییر کند، صفت متغیر نامیده می‌شود.

صفاتی را که بتوان برای آنها اندازه تعریف کرد، صفات متغیر کمی و صفاتی را که نتوان برای آنها اندازه‌ای تعریف کرد، صفات کیفی نامند.

صفات متغیر کمی یا گسسته (نا پیوسته) هستند یا پیوسته.

- **طبقه‌بندی صفات:** بسته به اینکه صفت متغیر، گسسته یا پیوسته باشد، آنها را با توجه به ماهیتشان طبقه‌بندی می‌کنند.

در حالت گسسته آنها را به ترتیب صعودی یا نزولی مرتب می‌کنند و توزیع صفت را بر حسب فراوانی مطلق بدست می‌آورند.

در حالت پیوسته بودن صفت، آنها را در فاصله طبقه‌بندی کرده و توزیع صفت متغیر را بر حسب فراوانی مطلق بدست می‌آورند.

- **نمودارها**: برای بیان چگونگی توزیع صفت متغیر، می‌توان از نمودارها استفاده کرد. برای صفات متغیر نا پیوسته از نمودارهای میله‌ای و چند گوش (پلی‌گون) استفاده می‌شود.

و برای صفات متغیر پیوسته از نمودار هیستوگرام (بافت نگار) استفاده می‌شود.

و برای صفات کیفی از نمودارهای میله‌ای یا پلی‌گون و یا دایره‌ای استفاده می‌شود.

- **فراوانی تجمعی(انباشته):** فراوانی تجمعی مربوط به x_i برابر فراوانی تمامی مقادیر صفت x را گویند که از x_i تجاوز نکند:

$$F(x_i) = n(x \le x_i)$$

- **فراوانی تجمعی نسبی:** فراوانی تجمعی نسبی از جمع فراوانیهای نسبی صفت x تا نقطه x_i را فراوانی تجمعی نسبی در نقطه x_i نامند:

$$F'(x_i) = f(x \le x_i) = \frac{n(x \le x_i)}{n}$$

- مفهوم اندازه‌گیری و مقیاس آنها

در اجرای یک تحقیق آماری، اولین مرحله، مشاهده داده‌های آماری می‌باشد و آنهم اندازه‌گیری صفات در جامعه است.

«نسبت دادن اعداد به عناصر یا افراد جامعه بر طبق قواعد معینی، اندازه‌گیری نامیده می‌شود»

– انواع مقیاس اندازه‌گیری

الف- مقیاس اسمی: در این نوع اندازه‌گیری، از اعداد به منظور جدا کردن نمودها یا عناصر طبقات مختلف استفاده می‌شود. این اعداد نقش علامت یا نام را برای طبقات دارند و هیچ معنای کمی ندارند و روی آنها هیچگونه عملیات ریاضی انجام نمی‌گیرد.

ب- مقیاس ترتیبی: دراین نوع اندازه‌گیری از اعداد برای مقایسه عناصر از نظر کوچکتر یا بزرگتر بودن استفاده می‌شود. در این مقیاس، عناصر بر اساس اندازه‌های نسبی مرتب می‌شوند. روی اعداد مقیاس ترتیبی نیز هیچگونه عملیات ریاضی انجام نمی‌گیرد.

ج- مقیاس فاصله‌ای: اگر عناصر مورد اندازه‌گیری در جامعه، نه تنها از نظر ترتیب بلکه بر حسب طول فاصله بین اندازه‌های عناصر نیز بتواند طبقه‌بندی شوند، آنگاه این مقیاس فاصله‌ای خواهد بود. مانند اندازه‌گیری شدت حرارت که با صفر مقایسه می‌شود.

د- مقیاس نسبی: وقتی اندازه‌گیری به منظور مقایسه نسبت خاصیت عناصر مورد مطالعه، در نظر گرفته شود، بطوریکه بتوان گفت که اندازه عنصر اول چند برابر عنصر دوم و چند برابر عنصر سوم و...است، این اندازه‌گیری با مقیاس نسبتی خواهد بود. در مقیاس نسبتی عملیات ریاضی انجام پذیر است.

– توزیع صفت متغیر (توزیع فراوانی صفات): اگر صفت متغیر، ناپیوسته بوده و دامنه تغییرات آن وسیع نباشد، آن را به فاصله یکی طبقه‌بندی کرده و بر حسب فراوانی مطلق (n_i) یا فراوانی نسبی (f_i) نشان می‌دهند:

x	x_1	x_2	x_3	x_s
n_i	n_1	n_2	n_3	n_s

و یا

x	x_1	x_2	x_3	x_s
f_i	f_1	f_2	f_3....	f_s

که در توزیع فوق f_i فراوانی نسبی نسبت متغیر x می‌باشد و از خارج قسمت فراوانی هر یک از متغیرها به حجم جامعه (n) بدست می‌آید:

$$f_i = \frac{n_i}{n}$$

همواره مجموع فراوانیهای نسبی مساوی یک می‌باشد.

در صورتیکه صفت متغیر ناپیوسته بوده و دامنه تغییرات آن وسیع باشد و یا صفت متغیر پیوسته باشد، آنرا در فاصله، طبقه‌بندی می‌کنند:

x	$x_1 - x_2$	$x_2 - x_3$	$x_{s-}\ x_{s+1}$
n_i	n_1	n_2	n_s

و یا

x	$x_1 - x_2$	$x_2 - x_3$	.....	$x_s - x_{s+1}$
f_i	f_1	f_2	.....	f_s

۲- مشخص‌کننده‌های عددی(آماری)

مشخص‌کننده‌های عددی، کمیتهائی(پارامترهائی) هستند که برای مقایسه بین چند جامعه بکار می‌روند.

و در سه گروه زیر طبقه‌بندی می‌شوند:

۲-۱- مشخص‌کننده‌های مرکزی: که مرکز تمرکز توزیع صفت را بیان می‌کنند و از واحد اندازه‌گیری صفت متغیر تبعیت دارند. مهمترین آنها عبارتند از:

– **میانگین حسابی یا متوسط حسابی:** که از مجموع متغیرها به حجم جامعه به دست می‌آید:

$$\overline{x} = \frac{\sum x_i}{n}$$

در صورتیکه جدول بر حسب فراوانی مطلق یا فراوانی نسبی باشد، میانگین حسابی بصورت:

$$\overline{x} = \frac{\sum n_i x_i}{n} \quad \text{یا} \quad \overline{x} = \sum f_i x_i$$

محاسبه میشود که به میانگین وزنی معروفند. که دراین رابطه $f_i = \frac{n_i}{n}$ فراوانی نسبی صفت متغیر می‌باشد.

– **خواص میانگین:**

۱- میانگین اعداد ثابت، خود عدد ثابت است: $\frac{\sum a}{n} = a$

۲- اگر از متغیرها، عدد ثابتی مانند a را کم یا اضافه کنیم از میانگین هم به اندازهa کم یا اضافه خواهد شد.

$$\frac{\sum (x_i \pm a)}{n} = \overline{x} \pm a$$

۳- اگرمتغیرها را در عدد ثابتی مانند k ضرب یا تقسیم کنیم، میانگین هم در عدد ثابت k ضرب و یا بر عدد ثابت k تقسیم می‌شود:

$$\frac{\sum (kx_i)}{n} = k\overline{x} \qquad \frac{\sum \left(\frac{x_i}{k}\right)}{n} = \frac{\overline{x}}{k}$$

۴- مجموع انحرافات متغیرها از میانگین مساوی صفر است:

$$\sum (x_i - \overline{x}) = ۰ \qquad \sum n_i (x_i - \overline{x}) = ۰ \qquad \sum f_i (x_i - \overline{x}) = ۰$$

– **میانه:** متغیری است که جامعه را به دو قسمت مساوی تقسیم می‌کند و بر اساس خاصیت فراوانی تجمعی محاسبه میشود:

$$F(x_i) < \frac{n}{۲} < F(x_{i+۱})$$

اگر متغیر x ناپیوسته باشد، متغیر متعلق به $F(x_{i+۱})$ میانه توزیع خواهد بود.

اگرمتغیر x پیوسته باشد، مقدار میانه از رابطه تقریبی زیر بدست می‌آید:

$$Me = x_i + \frac{n/۲ - F(x_i)}{n_i} \times I$$

که در این رابطه، x_i حد پائین طبقه‌ای از متغیر x است که میانه درداخل آن قرار دارد و به توسط $F(x_{i+۱})$ این طبقه مشخص می‌شود و I فاصله طبقه میانه است.

یکی از مهمترین خواص میانه، اینست که مجموع قدر مطلق انحرافات متغیرها از میانه کمترین مقدار ممکن است:

$$\sum |X_i - Me| = \min \quad \text{یا} \quad \sum n_i |x_i - Me| = \min$$

– **چارکها:** چارکها جامعه را به چهار قسمت مساوی تقسیم می‌کنند، چارکها نیز مانند میانه محاسبه می‌شوند. چارک اول که $1/۴$ حجم جامعه را مشخص می‌کند، بصورت:

$$Q_۱ = X_i + \frac{n/۴ - F(x_i)}{n_i} \times I$$

و چارک سوم که $\frac{۳}{۴}$ حجم جامعه را مشخص می‌کند بصورت:

$$Q_۳ = X_i + \frac{{}^{۳n}\!/_{۴} - F(x_i)}{n_i} \times I$$

محاسبه می‌شود.

– **نما یا مُد:** متغیری که دارای بزرگترین فراوانی(مطلق یا نسبی) در جامعه باشد، مد توزیع نامیده می‌شود. برای توزیع ناپیوسته، با توجه به جدول توزیع صفت متغیر، مُد توزیع به توسط بزرگترین فراوانی بدست می‌آید. برای توزیعهای پیوسته با استفاده از مفهوم چگالی، مقدار نما محاسبه می‌شود.

$$MO = X_i + \frac{d_m - d_{m-۱}}{۲d_m - d_{m-۱} - d_{m+۱}} \times I$$

که در این رابطه d چگالی فراوانی می‌باشد که از $d = \frac{n_i}{\Delta x}$ بدست می‌آید.

در صورتیکه فاصله طبقات یکسان باشد، چگالی ماکزیمم و فراوانی ماکزیمم بر یکدیگر منطبق می‌گردند و در این صورت اندازه نما بر حسب فراوانی مطلق یا نسبی بصورت زیر محاسبه می‌شود:

$$MO = X_i + \frac{n_m - n_{m-۱}}{۲n_m - n_{m-۱} - n_{m+۱}} \times I$$

همواره یکی از روابط زیر بین سه مشخص کننده مرکزی برقرار است:

توزیع قرینه است(متقارن است): $\bar{x} = Me = Mo$

توزیع دارای چولگی (انحراف از قرینگی) راست می‌باشد: $\bar{x} > Me > Mo$

توزیع دارای چولگی چپ می‌باشد: $\bar{x} < Me < Mo$

– **میانگین هندسی:** اگر متغیرها بصورت نسبت یا درصد و یا شاخص بیان شده باشند، در این صورت، متوسط متغیرها بصورت میانگین هندسی محاسبه می‌شود.

$$\bar{x}_G = \sqrt[n]{x_۱, x_۲, x_۳ \ldots x_n} \qquad (x_i > ۰)$$

– **میانگین همساز (هارمونیک):** اگر صفت متغیر درارتباط با زمان باشد و از دو واحد اندازه‌گیری تبعیت کند، دراین صورت برای محاسبه میانگین صفت متغیر، از میانگین همساز استفاده می‌شود:

$$\bar{x}_H = \frac{n}{\sum \frac{۱}{x_i}} \qquad \text{یا} \qquad \bar{x}_H = \frac{n}{\sum n_i \frac{۱}{x_i}}$$

همواره بین میانگین حسابی، میانگین هندسی و میانگین هارمونیک رابطهٔ زیر برقرار است:

$$\bar{x}_H < \bar{x}_G < \bar{x}_M$$

۲-۲- مشخص کننده‌های پراکندگی

– **دامنه تغییرات یا فاصله تغییرات** (Range): اختلاف بزرگترین متغیر و کوچکترین متغیر را دامنه تغییرات نامند:

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

– **انحراف چارکی(نیم دامنه چارکی):** برای اندازه پراکندگی توزیعهای باز(توزیعهائی که حد پایین یا حد بالای آنها نامعلوم است). محاسبه می‌شود:

$$Q = \frac{Q_۳ - Q_۱}{۲}$$

– **میانگین قدر مطلق انحرافات (انحراف متوسط):** از مجموع قدر مطلق اختلاف متغیرها از میانگین تقسیم بر n (حجم جامعه) بدست می‌آید:

$$M.D=\frac{\sum|x_i-\bar{x}|}{n}$$ یا $$M.D=\frac{\sum n_i|x_i-\bar{x}|}{n}$$

– **واریانس:** میانگین مجذور انحرافات متغیرها از میانگین را واریانس می‌نامند:

$$D(x)=\sigma^۲=\frac{\sum(x_i-\bar{x})^۲}{n}\quad\frac{\sum x_i^۲}{n}-\bar{x}^۲$$

در صورتیکه توزیع صفت متغیر بر حسب فراوانی بدست آمده باشد، واریانس مساوی خواهد شد با:

$$D(x)=\sigma^۲=\frac{\sum n_i(x_i-\bar{x})^۲}{n}=\frac{\sum n_i x_i^۲}{n}-\bar{x}^۲$$

واریانس با مجذور واحد اندازه‌گیری صفت متغیر بیان می‌شود، بهمین جهت از آن جذر گرفته و به‌عنوان معیار پراکندگی مناسب بدست می‌آورند.

– **انحراف معیار(انحراف استاندارد):** از جذر واریانس بدست می‌آید و هم بعد صفت متغیر می‌باشد:

$$\sigma=+\sqrt{D(x)}$$

– **خواص واریانس:**

۱- واریانس عدد ثابت مساوی با صفر است. $D(a)=۰$

۲- اگر از متغیرها عدد ثابتی مانند α را کم و یا به متغیرها عدد ثابتی مانند a را اضافه کنیم، در مقدار واریانس تأثیری نخواهد داشت: $D(X\pm a)=D(x)$

۳- اگر متغیرها رادر عدد ثابتی مانند k ضرب کنیم، واریانس در $k^۲$ ضرب خواهد شد: $D(kx)=k^۲D(X)$

۴- اگر متغیرها را بر عدد ثابتی مانند K تقسیم کنیم، واریانس بر $k^۲$ تقسیم خواهد شد:

$$D\left(\frac{X}{K}\right)=\frac{۱}{K^۲}D(X)$$

۲-۳- مشخص‌کننده‌های نسبی پراکندگی

– **ضریب تغییرات:** از خارج قسمت انحراف معیار (σ) به میانگین $(\bar{X})$ بدست می‌آید و از واحد اندازه‌گیری مجرد است:

$$C.V=\frac{\sigma}{\bar{X}}\times۱۰۰$$

ضریب تغییرات، درصد تغییرات متغیرها را حول میانگین نشان می‌دهد.

– **ضریب چولگی:** با استفاده از روابط بین سه مشخص کننده مرکزی، ضریب چولگی پیرسن محاسبه می‌شود که از دقت بالائی برخوردار نیست:

$$A_s=\frac{\bar{X}-MO}{\sigma}$$

در صورتیکه متغیر دارای چولگی ملایم باشد، همواره رابطه زیر بین سه مشخصه مرکزی برقرار است:

$$\bar{X}-MO=۳(\bar{X}-Me)$$

ولی برای محاسبه دقیق‌تر، ضریب چولگی از خارج قسمت گشتاور مرکزی مرتبه سوم $(\mu_۳)$ به $\sigma^۳$ استفاده می‌شود:

$$A=\frac{\mu_۳}{\sigma^۳}$$

که درآن $\mu_۳$بصورت زیر محاسبه می‌شود:

$$\mu_۳=\frac{\sum(x_i-\bar{x})^۳}{n}$$ یا $$\mu_۳=\frac{\sum n_i(x_i-\bar{x})^۳}{n}$$

ضریب چولگی ممکن است مثبت و یا منفی بدست آید. اگر مثبت باشد، توزیع دارای چولگی راست و در صورتی که منفی باشد، توزیع دارای چولگی چپ می‌باشد. اگر $A=0$ گردد، توزیع متقارن است. اگر $|A|<0/1$ باشد، توزیع تقریباً قرینه است و اگر $0/1\le|A|\le 0/5$ باشد، توزیع دارای چولگی ملایم و اگر $|A|>0/5$ باشد، توزیع شدیداً چولگی دارد.

گاهی اوقات ضریب چولگی را با استفاده از چارکها محاسبه می‌کنند:

$$A=\frac{Q_3-2Q_2+Q_1}{Q_3-Q_1}$$

- **ضریب کشیدگی:** از خارج قسمت گشتاور مرکزی مرتبه چهارم به σ^4 بدست می‌آید و معمولاً کشیدگی توزیعها را با توزیع نرمال مقایسه می‌کنند و از طرفی ضریب کشیدگی توزیع نرمال مساوی با ۳ می‌باشد، لذا ضریب کشیدگی بصورت:

$$K=\frac{\mu_4}{\sigma^4}-3$$

محاسبه می‌شود. اگر $K=0$ باشد، کشیدگی توزیع به اندازه نرمال است و اگر $K<0$ باشد، توزیع از نرمال کوتاهتر است. در صورتیکه $K>0$ باشد، توزیع از توزیع نرمال کشیده‌تر است. گشتاور مرکزی مرتبه چهارم بصورت زیر بدست می‌آید:

$$\mu_4=\frac{\sum(x_i-\bar{x})^4}{n} \qquad \text{یا} \qquad \mu_4=\frac{\sum n_i(x_i-\bar{x})^4}{n}$$

- **گشتاورها:** گشتاورها، میانگین‌هائی هستند حول مرکز ثقل توزیع با مراتب مختلف.

چون مرکز ثقل توزیع می‌تواند مبدأ مختصات(صفر) و یا میانگین حسابی باشد، لذا دو نوع گشتاور دراینجا تعریف می‌کنیم:

۱- گشتاورهای اولیه: یا گشتاورهای نسبت به‌مبدأ صفر بصورت زیر محاسبه می‌شوند:

$$m_h=\frac{\sum x^h}{n} \qquad \text{یا} \qquad m_h=\frac{\sum n_i x_i^h}{n} \qquad \text{یا} \qquad m_h=\sum f_i x_i^h$$

و گشتاورهای مرتبه اول تا چهارم نسبت به‌مبدأ صفر مساوی است با:

$$m_1=\frac{\sum x_i}{n}=\bar{x} \qquad m_2=\frac{\sum x^2}{n}=\bar{x}^2$$

$$m_3=\frac{\sum x_i^3}{n}=\bar{x}^3 \qquad m_4=\frac{\sum x_i^4}{n}=\bar{x}^4$$

که به گشتاورهای توانی یا میانگینهای توانی معروفند.

۲- گشتاورهای مرکزی: یا گشتاورهای نسبت به‌مبدأ میانگین که به صورت زیر محاسبه می‌شوند:

$$\mu_h=\frac{\sum(x_i-\bar{x})^h}{n} \qquad \text{یا} \qquad \mu_h=\frac{\sum n_i(x_i-\bar{x})^h}{n} \qquad \text{یا} \qquad \mu_h=\sum f_i(x_i-\bar{x})^h$$

و گشتاورهای مرتبه اول تا چهارم نسبت به مبدأ میانگین بصورت زیر بدست می‌آیند:

$$\mu_1=\frac{\sum(x-\bar{x})}{n}=0$$

$$\mu_2=\frac{\sum(x_i-\bar{x})^2}{n}=D(x)$$

$$\mu_3=\frac{\sum(x_i-\bar{x})^3}{n} \qquad , \qquad \mu_4=\frac{\sum(x_i-\bar{x})^4}{n}$$

- رابطه تبدیل گشتاورهای اولیه به گشتاورهای مرکزی: برای تبدیل گشتاورهای اولیه به گشتاورهای مرکزی ثابت می‌شود که $\mu_h=\frac{\sum(x_i-\bar{x})^h}{n}=(m-m_1)^h$ و در نتیجه برای $h=1,2,3,4$ روابط زیر بدست می‌آید:

$$\mu_1 = (m_\circ - m_1) = \circ$$
$$\mu_2 = (m - m_1)^2 = m_2 - m_1^2$$
$$\mu_3 = (m - m_1)^3 = m_3 - 3m_2 m_1 + 2m_1^3$$
$$\mu_4 = (m - m_1)^4 = m_4 - 4m_3 m_1 + 6m_2 m_1^2 - 3m_1^4$$

- مشخص کننده‌های عددی k جامعه: گاهی اوقات از ترکیب چند جامعه، جامعه کل بوجود می‌آید که در آنها صفت متغیر x دارای واحدهای اندازه‌گیری یکسان بوده ولی مشخص کننده‌های عددی در هر یک از جامعه‌های جزء متفاوت می‌باشد، لذا قانون توزیع صفت متغیر x در جامعه کل از ترکیب توزیع صفت x در جامعه‌های جزء بدست می‌آید ومی‌توان مشخص‌کننده‌های عددی را برای آنها محاسبه کرد. اگر صفت متغیر x در جامعه اول، دوم،...، k ام به ترتیب $x_1, x_2, \ldots, x_k$ در نظر گرفته شود و میانگین و واریانس برای هر یک از آنها وجود داشته باشد:

$$\begin{array}{llll} x_1: & n_1, & \mu_1, & \sigma_1 \\ x_2: & n_2, & \mu_2, & \sigma_2^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_k: & n_k, & \mu_k, & \sigma_k^2 \end{array}$$

آنگاه میانگین k جامعه مساوی خواهد شد با:

$$\bar{\mu} = \frac{\sum n_i \mu_i}{n}$$

و واریانس k جامعه مساوی خواهد شد با:

$$D(x) = \frac{\sum n_i \sigma_i^2}{n} + \frac{\sum n_i (\mu_i - \bar{\mu})^2}{n}$$

که درآنها $n = n_1 + n_2 + \ldots + n_k$ می‌باشد.

۳-توزیعهای دو بعدی

– دستگاه دو صفت متغیر: نتایج مشاهدات بر روی دو صفت متغیر y,x با هم مجموعه‌ای از زوجهای مرتب شده را تشکیل می‌دهند که دنباله زوجهای مرتب شده را سری همبستگی می‌نامند.

اگر توزیع فراوانیهای دو صفت را با هم در جدول نشان دهیم، جدول زیر بدست می‌آید:

X \ Y	y_1	y_2	y_t	$n_{i\circ}$
x_1	n_{11}	n_{12}	n_{1t}	$n_{1\circ}$
x_2	n_{21}	n_{22}	n_{2t}	$n_{2\circ}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
x_s	n_{s1}	n_{s2}	n_{st}	$n_{s\circ}$
$n_{\circ j}$	$n_{\circ 1}$	$n_{\circ 2}$	$n_{\circ t}$	n

اگر فراوانیهای توزیعهای دو بعدی را سطری جمع کنیم توزیع حاشیه‌ای صفت x و در صورتیکه فراوانیها را ستونی با یکدیگر جمع کنیم، توزیع حاشیه‌ای برای صفت y بدست می‌آید.

– **توزیعهای شرطی:** چون به تعداد سطرها توزیع تکی برای صفت y و به تعداد ستونها، توزیع تکی برای صفت x در توزیعهای دو بعدی وجود دارد و هر یک از آنها به ازاء مقدار معینی از صفت x و با مقدار معینی از صفت y بوجود می‌آیند، آنها را توزیعهای شرطی می‌نامند.

– **مفهوم بستگی آماری:** رابطه بین صفت متغیر که از افزایش یا کاهش یکی از صفت‌ها و تغییرات توزیعهای شرطی صفت دیگر بوجودمی‌آید، بستگی آماری نامیده می‌شود. در نتیجه هرگاه با تغییر یکی از صفتهای متغیر، توزیعهای شرطی صفت متغیر دیگر تغییر کند، گویند بستگی آماری بین دو صفت وجود دارد.

– **همبستگی:** هرگاه به ازاء افزایش یا کاهش یکی از صفت‌های متغیر، میانگین توزیعهای شرطی صفت متغیر دیگر تغییر کند، بین دو صفت متغیر x و y همبستگی وجود خواهد داشت.

چون در مفهوم همبستگی، دو صفت x و y هم ارز نمی‌باشند، لذا:

الف- همبستگی صفت y بر حسب x را رگرسیون y روی x می‌نامند و بصورت $\overline{y}_x = f(x)$ نوشته می‌شود.

ب- همبستگی صفتx بر حسب y را رگرسیون xروی y می‌نامند و بصورت $\overline{x}_y = g(y)$ می‌نویسند. بنابراین اگر با تغییر صفت x، میانگین توزیعهای شرطی y $(\overline{y}_i)$ تغییر کنند، همبستگی y روی x وجود دارد و هم چنین اگر با تغییر صفت y میانگین توزیعهای شرطی x $(\overline{x}_j)$تغییر کنند، دلیل بر وجود همبستگی یا رگرسیون x روی y خواهد بود.

اگر زوجهای مرتب شده $(x_i,\overline{y}_i)$ و یا $(\overline{x}_j,y_j)$رادر صفحه مختصات مشخص کنیم، از طرز قرار گرفتن نقاط در صفحه مختصات، نوع معادله‌ای که بیان کننده رابطه بین دو صفت(نوع معادله رگرسیون) می‌باشد، بدست می‌آید.

در صورتیکه رابطه رگرسیون به صورت معادله خطی مستقیم باشد، یعنی:

$\overline{y}_x = f(x) = a + bx$ یا $\overline{x}_y = g(y) = a' + b'y$

درآن صورت ضرایب آن با استفاده از روش کمترین توانهای دوم بدست می‌آید:

$$a = \overline{y} - b\overline{x}$$

$$b = \frac{sp_{xy}}{ss_x} = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(x-\overline{x})^2} = \frac{\sum xy - \frac{\sum x.\sum y}{n}}{\sum x^2 - \frac{\left(\sum x\right)^2}{n}} = \frac{cov(x,y)}{D(x)}$$

– **خطای معادله‌رگرسیون:** اگر معادله خط همبستگی(معادله خط رگرسیون) بصورت $\underset{x}{\bar{y}} = a + bx$ بیان شده باشد، آنگاه خطای معادله رگرسیون $\left(\sigma_{y.x}\right)$ از جذر واریانس همبستگی یا واریانس رگرسیون بدست می‌آید:

$$\sigma_{y.x}^{2} = \frac{\sum y^{2} - a\sum y - b\sum xy}{n}$$

$$\sigma_{y.x} = \sqrt{\frac{\sum y^{2} - a\sum y - b\sum xy}{n}}$$

از خطای معادله رگرسیون برای ارزیابی صفت y به ازاء یک مقدار معین از صفت x استفاده می‌کنند.

$$\underset{x_i}{\bar{y}} - \sigma_{y.x} < y < \bar{y}_{xi} + \sigma_{y.x}$$

– **ضریب همبستگی:** به منظور تعیین شدت همبستگی بین دو صفت x وy از ضریب همبستگی استفاده می‌شود که: اولاً ضریب همبستگی بعد ندارد یعنی از هیچ واحد اندازه‌گیری تبعیت نمی‌کند، ثانیاً حدود تغییرات ضریب همبستگی همواره بین $(-1, +1)$ تغییر می‌کند.

اندازه ضریب همبستگی به توسط یکی از روابط زیر بدست می‌آید.

$$\rho = \frac{\operatorname{cov}(x, y)}{\sigma_x . \sigma_y} = \frac{sp_{xy}}{\sqrt{ss_x . ss_y}} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^{2} . \sum(y - \bar{y})^{2}}} = \sqrt{b.b'}$$

اگر $\rho = 0$ باشد، همبستگی خطی بین دو صفت وجود ندارد.

اگر $\rho = +1$ باشد، همبستگی کامل و مستقیم است.

اگر $\rho = -1$ باشد، همبستگی کامل و معکوس است.

اگر $-1 < \rho < +1$ باشد، همبستگی را ضعیف و یا شدید می‌نامند.

– **ضریب تعیین یا ضریب تشخیص:** از خارج قسمت واریانس معادله رگرسیون $\left(\sigma_{y.x}^{2}\right)$ به واریانس y یعنی σ_y^{2} در معادله رگرسیون y بر حسب Xبدست می‌آید:

$$R^{2} = \frac{\sigma_{y.x}^{2}}{\sigma_y^{2}} = \frac{\left[c_{ov}(x, y)\right]^{2}}{\sigma_x^{2} . \sigma_y^{2}} = \rho^{2}$$

که در این رابطه $C_{ov}(x, y)$، کواریانس نامیده می‌شود که پراکندگی توأم دو متغیر xوy را نشان می‌دهد و به نام گشتاور همبستگی نیز موسوم است و بصورت زیر محاسبه می‌شود.

$$\operatorname{Cov}(x, y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{n} = \frac{\sum xy}{n} - \bar{x}\,\bar{y}$$

ضریب تعیین، میزان تغییر پذیری مقادیر مشاهده شده صفت y را حول معادله رگرسیون به میزان تغییرپذیری صفت x نشان می‌دهد. یعنی اگر ρ^{2} را در ۱۰۰ ضرب کنیم و بر حسب درصد بیان نمائیم، آنگاه R^{2} درصد تغییرات صفت y را ناشی از اثرات صفت متغیر x نشان می‌دهد.

تمرینات بخش اول

۱- اندازه صفت متغیر x در جامعه بصورت زیر بدست آمده است:

x : ۲۰ ,۱۸ ,۲۵ ,۱۹ ,۱۶ ,۱۴ ,۲۳

میانگین و میانه توزیع را محاسبه کنید.

۲- جدول زیر توزیع صفت متغیر x رانشان می‌دهد:

x_i	۴–۸	۸–۱۲	۱۲–۱۶	۱۶–۲۰	۲۰–۲۴	۲۴–۲۸
فراوانی	۵	۱۲	۱۸	۱۵	۲۰	۱۰

ضریب تغییرات x را محاسبه کنید.

۳- توزیع صفت متغیر x بصورت جدول زیر است:

x	–۵	۵–۱۰	۱۰–۱۵	۱۵–۲۰	۲۰–
فراوانی	۳	۸	۱۲	۴	۳

نیم دامنه چارکی توزیع را محاسبه کنید.

۴- اگردر جامعه‌ای میانگین $\bar{x}=۵۰$ و واریانس $D(x)=۶۴$ بدست آمده باشد، ضریب تغییرات y که بر طبق رابطه $y=۲x+۱۰$ از صفت x تبعیت می‌کند، چقدر است؟

۵- نرخ رشد سالیانه سود شرکت تجارتی A در پنج سال گذشته بصورت زیر بدست آمده است:

X : ۲/۹ ,۲/۸۵ ,۲/۴۵ ,۲/۱۵ ,۲/۱

متوسط نرخ رشد سود سالیانه چقدر است؟

۶- قیمت کالائی در سال گذشته ۲۰٪ کاهش و امسال ۲۰٪ افزایش داشته است، متوسط نرخ رشد قیمت این کالا دراین دو سال چقدر است؟

۷- فرض کنید شاخص قیمت خرده‌فروشی از ۲۰۰ درسال ۱۳۷۸ به ۴۵۰ درسال ۱۳۸۰ رسیده باشد، متوسط نرخ تورم سالانه دراین فاصله زمانی چقدر بوده است؟

۸- سه دوچرخه سوار، یک مسافت را بصورت زیر طی می‌کنند:

اولی با سرعت ۱۵ کیلومتر در ساعت، دومی با سرعت ۲۰ کیلومتر در ساعت و سومی با سرعت ۳۰ کیلومتر در ساعت . متوسط سرعت آنان چقدراست؟

۹- گشتاورهای مرتبه اول، دوم و سوم اولیه برای صفت متغیر بقرار زیر بدست آمده است.

$m_۱=۲ \quad m_۲=۸۵ \quad m_۳=۵۰$

ضریب چولگی توزیع را محاسبه کنید.

۱۰- زمان لازم برای ساختن واحد محصول برای ۳۲ کارگر در یک واحد صنعتی در یک شیفت کار بصورت جدول زیر بدست آمده است.

(زمان صرف شده برای ساخت واحد محصول) x_i	۱۰	۱۲	۱۲	۱۶	۱۸
(تعداد کارگران) n_i	۲	۳	۷	۱۲	۸

متوسط زمان لازم برای ساختن یک واحد محصول چقدر است؟

۱۱- توزیع دو صفت متغیر x و y بصورت جدول زیر بدست آمده است:

x \ y	۱	۲	۳	۴	جمع
۰	۰	۱	۳	۱	۵
۲	۱	۲	۳	۲	۸
۴	۱	۳	۳	۱	۸
۶	۱	۲	۱	۰	۴
جمع	۳	۸	۱۰	۴	۲۵

کواریانس x وy را محاسبه کنید.

۱۲- توزیع دو صفت متغیر x وy در جامعه بصورت جدول زیر است:

x_i	۱	۲	۳	۴	۵
y_i	۱۰	۸	۷	۵	۳

کواریانس یا گشتاور همبستگی بین دو صفت متغیر x وy چقدراست؟

۱۳- در جامعه‌ای دو بعدی پس از محاسبات لازم، کمیتهای زیر بدست آمده است:

$SpD_{xy} = ۷۰۶ \qquad SSD_x = ۶۲۱ \qquad \bar{X} = ۷۵ \qquad \bar{y} = ۱۲۵/۵$

معادله خط رگرسیون y روی X را بدست آورید.

۱۴- در یک جامعه نرمال دو بعدی $C_{ov}(x, y) = ۲۰$ ، $\sigma_x^۲ = ۲۵$ و $\sigma_y^۲ = ۳۶$ بدست آمده است. ضریب همبستگی بین دو صفت x و y را محاسبه کنید.

۱۵- در جامعه نرمال دو بعدی، کمیتهای زیر محاسبه شده‌اند:

$\sigma_y = ۳۶ \;,\; \sigma_x = ۲۵ \quad , R^۲ = ۰/۷۲$

ضریب زاویه خط رگرسیون y روی x کدام است؟

بخش دوم

نظریه احتمال

نظریه احتمال

– **مفاهیم اساسی:** نظریه احتمال بر پایه دو مفهوم حادثه و آزمایش استوار است و بعنوان مفاهیم اولیه بکار می‌روند، بهمین دلیل آنها را تعریف نکرده، بلکه بصورت زیر توصیف می‌کنند:

حادثه: هرنمود یا پدیده‌ای که درنتیجه آزمایش، وقوع یا عدم وقوع آن بتواند مورد بحث قرار گیرد، حادثه نامیده میشود.

آزمایش: برقرار کردن مجموعه شرطهای معین که بطور مکرر و به تعداد دلخواه زیادی عملی گردد، آزمایش نامند.

حادثه درنتیجه آزمایش به یکی از سه صورت زیر اتفاق می‌افتد:

۱- حادثه یقین ۲- حادثه غیر ممکن ۳- حادثه تصادفی

– **تعریف احتمال حادثه:** اندازه امکان وقوع هر حادثه را احتمال آن حادثه می‌نامند و با $P(A)$ نشان می‌دهند.

– **فضای نمونه:** مجموعه‌ای از عناصر دلخواه را فضای نمونه می‌نامندو این عناصر دلخواه، تمامی نتایج ممکن آزمایش هستند یا بعبارتی مجموعه تمامی حوادث ساده را در یک آزمایش فضای نمونه می‌نامند.

اندازه امکان وقوع حادثه به یکی از سه صورت زیر محاسبه می‌گردد.

– محاسبه احتمال بر اساس تعریف کلاسیک احتمال: تعریف کلاسیک احتمال بر پایه هم احتمال بودن حوادث ویا هم امکان بودن حوادث بصورت زیر بیان می‌شود:

احتمال حادثه A از خارج قسمت تعداد حالات مساعد بر حادثه A به تعداد تمامی نتایج ممکن آزمایش بدست می‌آید:

$$P(A)=\frac{\text{تعداد حالات مساعد بر حادثه}A}{\text{تعداد تمامی نتایج ممکن}}$$

– محاسبه احتمال بر اساس تعریف آماری احتمال- تعریف آماری احتمال برای حوادثی که هم احتمال نباشند بکار می‌رود و از خاصیت فراوانی نسبی حادثه استفاده می‌شود، زیرا فراوانی نسبی را فقط می‌توان پس از انجام آزمایش‌ها تعیین کرد، بنابراین در تکرار زیاد آزمایشها، فراوانی نسبی حادثه به سمت عددی میل می‌کند که آن عدد، اندازه امکان وقوع حادثه را بیان می‌کند و احتمال آن حادثه نامیده می‌شود:

$$P(A)=\underset{n\to\infty}{f}(A)=\frac{m}{n}$$

این تعریف بر اساس قانون اعداد بزرگ ضعیف بصورت برنولی، بیان می‌شود:

$$\lim_{n\to\infty} p\left(\left|\frac{m}{n}-p\right|<\varepsilon\right)=۱$$

یعنی یقیناً در تکرار زیاد آزمایش‌ها، فراوانی نسبی حادثه به سمت احتمال حادثه میل می‌کند.

و بر همین اساس قانون اعداد بزرگ قوی بر نولی بصورت

$$P\left(\lim_{n\to\infty}\frac{m}{n}=P\right)=۱$$

تعریف می‌شود، یعنی هرگاه حد فراوانی نسبی حادثه برابر p گردد، یک حادثه یقین است.

– محاسبه احتمال بر اساس تعریف احتمال هندسی: اگر تعداد وقوع حوادث در نتیجه آزمایش شمارش ناپذیر باشد، در این صورت خارج قسمت اندازه امکان وقوع حادثه مورد نظر به اندازه امکان نتایج کل آزمایش را احتمال هندسی می‌نامند.

$$P(A)=\frac{mes(A)}{mes(S)}=\frac{(A)\ \text{اندازه امکان وقوع حادثه}}{(S)\ \text{اندازه امکان وقوع کل حادثه}}$$

- اصول متعارف احتمال

۱- احتمال حادثه بین صفر تا ۱ تغییر می‌کند:

$$0\le P(A)\le 1$$

۲- مجموع احتمالها در فضای نمونه(گروه کامل حوادث) مساوی یک است:

$$\sum_{i=1}^{n}P(A_i)=1$$

۳- اگر دو حادثه A و B نسبت به یکدیگر ناسازگار باشند، در این صورت، احتمال اجتماع دو حادثه مساوی است با اجتماع احتمال آن دو حادثه:

$$P(A\cup B)=P(A)+P(B)$$

۲-۱ قضایای احتمالات

- قضیه عمومی حاصل ضرب احتمالها: اگر حادثه A قابل تجزیه به حوادث جزئی مانند B و C باشد، بطوریکه حوادث B و C با هم بتوانند اتفاق بیفتند، در این صورت آنرا بصورت حاصلضرب دو حادثه می‌نویسند:

$$A=BC=B\cap C$$

و احتمال حادثه A مساوی خواهد شد با:

$$P(A)=P(B\cap C)=P(B).P(C|B)$$

و $P(C|B)$ را احتمال شرطی حادثه C بر حسب B می‌نامند و این در حالت دو حادثه B و C نسبت به یکدیگر وابسته هستند و درحالیکه دو حادثه B و C نسبت به یکدیگر مستقل باشند. احتمال حادثه A مساوی خواهد شد با:

$$P(A)=P(B\cap C)=P(B).P(C)$$

قضیه فوق قابل تعمیم برای بیش از دو حادثه است.

- قضیه عمومی حاصل جمع احتمالها: اگر حادثه A قابل تجزیه به حوادث جزء مانند Bو C باشد، بطوریکه در نتیجه آزمایش، حداقل یکی از حوادث(B یاC) بتواند اتفاق بیفتد. درآن صورت حادثه A از اجتماع حوادث جزء بدست می‌آید و احتمال حادثه A برابر است با:

$$A=B+C=B\cup C$$

$$P(A)=P(B)+P(C)-P(BC)$$

در این حالت دو حادثه B و C نسبت به یکدیگر سازگارند، ولی در صورت ناسازگار بودن حوادث B و C نسبت به یکدیگر، احتمال حادثه A از اجتماع احتمالهای حوادث B و C بدست می‌آید:

$$P(A)=P(B)+P(C)$$

قضیه فوق قابل تعمیم برای بیش از دو حادثه می‌باشد.

- قضیه احتمال متوسط: اگر حادثه H قابل تجزیه به حوادث جزء مانند $H_1,H_2.......,H_s$ باشد، که دو به دو نسبت به یکدیگر ناسازگارند و حادثه A بتواند با هر یک ازحوادث H_i اتفاق بیفتد، درآن صورت:

$$P(A)=\sum P(H_i).P(A|H_i)$$

- قضیه احتمال فرضیه‌ها(قضیه بیس): چنانچه در قضیه بالا، درنتیجه آزمایش، حادثه A وقوع کرده باشد، احتمال اینکه این حادثه برای فرضیه، iام (H_i) اتفاق افتاده باشد، مساوی است با احتمال فرضیه i ام در احتمال حادثه A بشرط فرضیه iام تقسیم بر احتمال متوسط:

$$P(H_i|A)=\frac{P(H_i).P(A|H_i)}{P(A)}\quad\frac{P(H_i).P(A|H_i)}{\sum P(H_i).P(A|H_i)}$$

– **آزمایشهای تکراری (فرمول برنولی):** اگر احتمال حادثه همواره ثابت باشد و برابر $P(A)=P$ و آزمایش را n بار تکرار کنیم، احتمال اینکه درn آزمایش مستقل از هم، mبار حادثه مورد نظر وقوع کند، از فرمول برنولی بدست می‌آید:

$$P_n(m)=\binom{n}{m}P^m\ q^{n-m} \qquad m=۰,۱,۲,.....,n$$

که درآن q احتمال عدم وقوع حادثه است.

$$p(\bar{A})=۱-P=q$$

– **فرمول مجانبی لاپلاس:** در صورتی که n عددنسبتاً بزرگی باشد، محاسبه آن بر اساس فرمول برنولی مشکل خواهد بود، ولی بر اساس قضایای حدی، از فرمول مجانبی لاپلاس می‌توان احتمال حادثه را بدست آورد، زیر ثابت می‌شود که با افزایش n، این احتمال $\underset{n}{p}(m)$ تقریباً بر طبق قانون نرمال توزیع می‌شود:

$$P_n(m)=\frac{۱}{\sqrt{npq}}\varphi(u)=\frac{۱}{\sqrt{npq}},\frac{۱}{\sqrt{۲\pi}}e^{-\frac{u^۲}{۲}}$$

– **قضیه انتگرالی لاپلاس– موآور:** اگر احتمال حادثه همواره ثابت و برابر P باشد و آزمایش را nبار تکرار کنیم، احتمال اینکه تعداد وقوع حادثه (m) بین $m_۱$ تا $m_۲$ بار وقوع کند، مساوی است با تابع لاپلاس به ازاء $U_۲$ منهای تابع لاپلاس به ازاء $U_۱$ یعنی:

$$P(m_۱\le m\le m_۲)=\phi(U_۲)-\phi(U_۱)$$

که دراین رابطه $U_۱$و $U_۲$ متغیرهای استاندارد شده $m_۱$ و $m_۲$ هستند.

$$U_۱=\frac{m_۱-np}{\sqrt{npq}} \qquad U_۲=\frac{m_۲-np}{\sqrt{npq}}$$

و تابع لاپلاس بصورت زیر بدست می‌آیند.

$$\phi(U)=\int_۰^{u_۰}\frac{۱}{\sqrt{۲\pi}}\ e^{-\frac{u^۲}{۲}}\ du$$

که در جدولی به همین نام به ازاء مقادیر مختلف u محاسبه شده است و با استفاده از آن، احتمال حادثه بدست می‌آید.

تمرینات بخش دوم

۱- از جعبه‌ای محتوی ۵ مهره سفید، ۳ مهره قرمز، ۲ مهره بطور تصادفی انتخاب می‌کنیم احتمال اینکه هر دو مهره هم رنگ باشند، چقدراست؟

۲- یک مؤسسه تولیدی دارای ۲۰ نفر پرسنل است که ۱۲نفر آنها مرد و بقیه زن هستند از این مؤسسه دو نفر را بطور تصادفی انتخاب می‌کنیم، احتمال اینکه یک مرد و یک زن انتخاب شود، چقدراست؟

۳- احتمال اصابت تیر بهدف ۰/۹ است. به هدف دوبار تیراندازی می‌شود، احتمال اینکه حداقل یک تیر بهدف اصابت کند، چیست؟

۴- از لوازم یدکی موجود درانبار ۳۰٪ متعلق به کارخانه A ، ۲۰٪ متعلق به کارخانهB و ۵۰٪ متعلق به کارخانه C است. احتمال خرابی لوازم یدکی برای کارخانهA، ۰/۰۵ برای کارخانه B ، ۰/۰۴ و برای کارخانه C ۰/۰۶ است اگر یک قطعه به تصادف انتخاب کنیم احتمال خراب بودن آن چقدراست؟

۵- احتمال اینکه در ۳ بار تیراندازی، حداقل یک تیر بهدف اصابت کند، مساوی ۰/۸۷۵ است، احتمال اصابت تیر بهدف چقدراست؟

۶- احتمال اینکه محصول تولید شده، توسط یک دستگاه اتومات، استاندارد باشد ۰/۸ است، تعداد ۱۰۰ محصول را بطور تصادفی انتخاب می‌کنیم، احتمال اینکه تعداد محصولات استاندارد بین ۷۵ تا ۹۰ باشد، چقدراست؟

۷- اگر $P(A)=\frac{۱}{۳}$ و $P(B)=\frac{۱}{۵}$ و $P(A|B)=\frac{۱}{۳}$ باشد $P(A\cup B)$ را محاسبه کنید.