آنالیز ترکیبی

شمارش: فرض کنید مقاله A به سه طریق (۱، ۲ و ۳) و مقاله B به دو طریق (M و N) انجام شود. می‌خواهیم ترکیبی دوتایی از این دو مقاله (اعداد و حروف) داشته باشیم به طوری که هر مقاله با یک عدد شروع و با یک حرف پایان بپذیرد.

همه‌ی حالت‌های ممکن را می‌توان به صورت زیر نمایش داد: $\{1N, 1M, 2N, 2M, 3N, 3M\}$

حال اگر تعداد اعداد و حروف زیاد باشد، نمایش همه‌ی آن‌ها وقت‌گیر خواهد بود و درصد خطا بالا می‌رود. به همین دلیل نیازمند به استفاده از اصول و قواعد شمارش می‌باشیم.

حال به بررسی دو اصل جمع و اصل ضرب که دو اصل مربوط به شمارش هستند، می‌پردازیم:

اصل جمع: با توجه به مقاله A و B، فرض کنید مقاله C یکی از مقاله‌های A یا B باشد، ولی هر دو نباشد.

با توجه به مقاله‌های A و B می‌خواهیم بدانیم که به چند طریق می‌توانیم مقاله C را انجام دهیم. بدیهی است که $2+3=5$ حالت امکان‌پذیر است. بنابراین اگر مقاله A، n طریق و مقاله B، m طریق قابل اجرا باشد، مقاله C، $m+n$ طریق قابل انجام دادن است.

اصل ضرب: با در نظر گرفتن مقاله A و B که در بالا معرفی شده است، فرض کنید مقاله C، انجام مقاله A و B به طور متوالی و یکی پس از دیگری است. اصل ضرب بیان می‌کند که مقاله C به $n\times m$ طریق صورت می‌پذیرد.

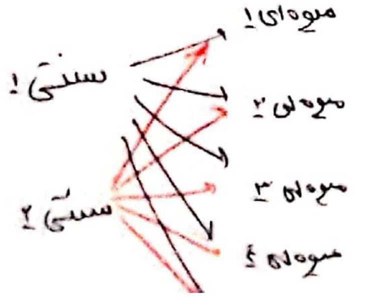
به عبارت دیگر، اگر انجام کاری شامل دو مرحله باشد، به طوری که برای انجام مرحله‌ی اول n روش و برای هر کدام از این n روش، مرحله‌ی دوم را بتوان به m روش انجام داد، در کل، کار مورد نظر با $m\times n$ روش قابل اجرا است.

مثال ۱: فرض کنید یک بستنی‌فروشی ۵ نوع بستنی میوه‌ای و ۲ نوع بستنی سنتی وجود دارد. برای خرید یک نوع بستنی، چند انتخاب وجود دارد؟

حل: طبق اصل جمع داریم:

$$5+2=7$$

مثال ۲: در مثال ۱، اگر انتخاب بستنی به صورت مخلوط یعنی ۱ نوع بستنی میوه‌ای، ۱ نوع بستنی سنتی مدنظر باشد، آن‌گاه تعداد انتخاب را محاسبه نمایید.



حل: طبق اصل ضرب داریم:

$$5\times 2=10$$

نکته: به صورت استقرایی می توان اصل جمع و اصل ضرب را برای k حالت نیز تعمیم داد.

تعمیم اصل جمع: اگر بتوان کاری را به k روش انجام داد، بطوریکه در روش اول m_1 انتخاب، در روش دوم m_2 انتخاب، ... ، و در روش k-ام، m_k انتخاب وجود داشته باشد و m_i ها دو به دو مجزا باشند، برای انجام کار مورد نظر

$$m_1+m_2+\dots+m_k=\sum_{i=1}^{k} m_i$$

روش وجود دارد.

تعمیم اصل ضرب: اگر انجام کاری شامل k مرحله باشد، بطوریکه برای انجام مرحله‌ی اول، m_1 روش، برای انجام مرحله دوم m_2 روش، ... و برای انجام مرحله‌ی k-ام، m_k روش وجود داشته باشد، و در هر مرحله انتخاب هر یک روش های آن ممکن باشد، آنگاه کار مورد نظر با $m_1\times m_2\times\dots\times m_k=\prod_{i=1}^{k} m_i$ روش قابل اجرا است.

مثال ۳: پنج دانشجوی سال دوم و ۴ دانشجوی سال اول می خواهند پینگ پنگ بازی کنند، به نحوی که هر کدام از دانشجویان سال دوم فقط و فقط یکبار با هر کدام از دانشجویان سال اول بازی کنند. تعداد کل بازی ها را محاسبه کنید.

حل: با توجه به اصل ضرب.

۴×۵=۲۰

مثال ۴: یک سکه و یک تاس را پرتاب کرده ایم. تعداد کل حالات ممکن چقدر است؟

حل: با توجه به اصل ضرب. ۲×۶=۱۲

مثال ۵: چند عدد ۲ رقمی زوج با ارقام متمایز وجود دارد؟

دهگان: {۱,۲,۳,...,۹} یکان: {۰,۲,۴,۶,۸}

حل: پس در جایگاه دهگان ۹ رقم، و در جایگاه یکان ۵ رقم می تواند قرار گیرد. که با توجه به اصل ضرب تعداد کل حالات می شود: ۹×۵=۴۵

ولی از آنجایی که گفته شده ارقام متمایز باشند، پس اعداد ۲۲، ۴۴، ۶۶ و ۸۸ را باید حذف کنیم.

بنابراین ۴۵-۴=۴۱ عدد دو رقمی زوج با ارقام متمایز می توانیم بنویسیم.

مثال ۶: مریم برای پر کردن اوقات فراغت خود قصد دارد به یکی از کلاس های هنری یا ورزشی برود. اگر او به کلاس هنری برود، می تواند در یکی از رشته های نقاشی یا خوشنویسی ثبت نام کند. ولی اگر قصد رفتن به کلاس ورزشی را داشته باشد، گزینه های او شنا، شطرنج و یا بسکتبال می باشد. الف: مریم چند انتخاب برای رفتن به کلاس هنری یا ورزشی دارد.

ب: 〃 〃 〃 〃 〃 〃 〃 و 〃 〃 .

حل: الف: طبق اصل جمع داریم: ۲+۳=۵ .

ب: 〃 〃 ضرب 〃 : ۲×۳=۶ .

مثال ۷ : شخصی می‌خواهد از شهر A به شهر C مسافرت کند و باید حتماً از شهر B عبور کند. اگر مسافرت از شهر A به شهر B چهار طریق و از شهر B به شهر C ۲ طریق قابل انجام باشد.

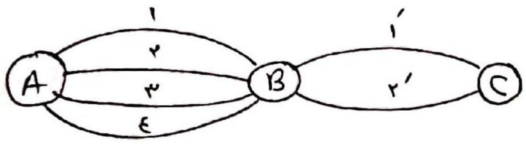
الف: مسافرت از شهر A به شهر C به چند طریق ممکن است؟

ب: چند مسیر وجود دارد که از A به C برود و به A برگردد؟

ج: " " " " " " " " " " ، طوریکه مسیر رفت و برگشت اصلاً اشتراکی نداشته باشد؟

د: " " " " " " " " " " ، رفت دقیقاً مثل مسیر برگشت نباشد؟

حل: الف: طبق اصل ضرب داریم $4\times 2=8$.



مسیرها عبارتند از $\{11', 12', 21', 22', 31', 32', 41', 42'\}$ می‌باشد.

ب: $8\times 8=64$

ج: $8\times 3=24$ (برای هر مسیر رفت، فقط سه مسیر برگشت وجود دارد که دقیقاً مثل مسیر رفت نباشد.)

د: $8\times 7=56$

مثال ۸ : چند عدد ۴ رقمی می‌توان ساخت، طوریکه:

الف: تکرار ارقام مجاز است.

ب: " " " " نباشد.

حل ۸:

هزارگان	صدگان	دهگان	یکان
۹	۱۰	۱۰	۱۰ حالت

هزارگان: $\{1,2,\dots,9\}$ — صدگان: $\{0,1,2,\dots,9\}$

الف: $9\times 10\times 10\times 10=9000$ با تکرار

ب: ۹ حالت ، ۹ حالت ، ۸ حالت ، ۷ حالت

بدون تکرار $9\times 9\times 8\times 7=4536$

مثال ۹ : چند عدد ۴ رقمی با ارقام زوج می‌توان ساخت؟

حل ۹: ارقام مجاز $=\{0,2,4,6,8\}$ (فرض با تکرار)

۴ ، ۵ ، ۵ ، ۵

$4\times 5\times 5\times 5=500$

مثال ۱۰ : چند عدد ۳ رقمی وجود دارد، که:

الف: عدد فرد باشد.

ب: عدد زوج باشد.

ج: عدد مضرب ۵ باشد.

د: عدد مضرب ۴ باشد.

($2k+1$ نمایش عدد فرد می‌باشد)

حل ۱۰: اعداد سه رقمی یعنی از ۱۰۰ تا ۹۹۹، پس:

الف:

$$99 < 2k+1 \le 999$$

طرفین منهای ۱

$$98 < 2k \le 998$$

طرفین تقسیم بر ۲

$$49 < k \le 499$$

$$\rightarrow k=499-49=\boxed{450}$$

ب: $100 \le 2k < 1000 \Rightarrow 50 \le k < 500 \Rightarrow k=500-50=\boxed{450}$

ج: $100 \le 5k < 1000 \Rightarrow 20 \le k < 200 \Rightarrow k=200-20=\boxed{180}$

د: $100 \le 4k < 1000 \Rightarrow 25 \le k < 250 \Rightarrow k=250-25=\boxed{225}$

* $a \le n \le b \Longrightarrow n=b-a+1$

$\begin{cases} a < n \le b \\ a \le n < b \end{cases} \Longrightarrow n=b-a$

$a < n < b \Longrightarrow n=b-a-1$

جایگشت: رمز هر کارت عابر بانک، چهار رقمی است. که از میان ارقام ۰ تا ۹ انتخاب می‌شود. طبق اصل ضرب، تعداد کل حالت‌های موجود، برای رمز کارت عابر بانک برابر با $10^4 = 10000$ می‌باشد. بنابراین 10000 حالت مختلف برای انتخاب رمز کارت وجود دارد. هر کارت نیز فقط یک رمز دارد، یعنی یکی از این اعداد چهار رقمی، به رمز هر کارت تعلق می‌گیرد.

تعریف: اگر چند شیء متمایز وجود داشته باشد، آن گاه، هر حالت قرار گرفتن آن‌ها کنار هم، یک جایگشت از آن اشیاء می‌گویند.

فاکتوریل: به ازای هر عدد حسابی $n \geq 0$، عبارت $n!$ خوانده می‌شود، n فاکتوریل و به صورت زیر تعریف می‌شود:

$$0! = 1 \quad , \quad 1! = 1 \quad , \quad 2! = 2\times1 = 2 \quad , \quad 3! = 3\times2\times1 = 6 \quad , \ldots ,$$

$$n! = n(n-1)(n-2)\ldots\times 3\times2\times1 \quad , \quad n \geq 1$$

مثال ۱۱: حاصل عبارت‌های زیر را به دست آورید.

$$\frac{8!}{5!} \overset{\text{حل}}{=} \frac{8\times7\times6\times5!}{5!} = 8\times7\times6 = \boxed{336} \quad , \quad \frac{n!}{(n-1)!} \overset{\text{حل}}{=} \boxed{n}$$

مثال ۱۲: به چند طریق می‌توان ۳ دانشجوی رشته کامپیوتر و ۴ دانشجوی رشته حسابداری را در کنار یکدیگر در یک صف مرتب کرد، به طوری که دانشجویان هم‌رشته در کنار یکدیگر باشند؟

(کامپیوتر کامپیوتر کامپیوتر) A　　(حسابداری حسابداری حسابداری حسابداری) B

حل: دانشجویان کامپیوتر، ۳! طریق جابه‌جایی دارند.

" حسابداری ۴! " " " و

گروه A و B هم می‌توانند با هم جابه‌جا شوند، ۲! طریق. پس تعداد کل حالت‌ها بنا به اصل ضرب می‌شود $2!\times4!\times3! = \boxed{288}$

مثال ۱۳: به چند طریق می‌توان ۲ مهره سفید و ۳ مهره سیاه را در یک خط مرتب کرد، به طوری که مهره‌های سفید در کنار یکدیگر باشند؟ (همه‌ی مهره‌ها متمایزند).

(سفید سفید) سیاه سیاه سیاه
1　2　3　4

حل: مهره‌های سفید، ۲! طریق جابه‌جا می‌شوند و در یک گروه قرار می‌دهیم.

حال چهار گروه می‌توانند به ۴! طریق جابه‌جا شوند. پس تعداد کل حالات بنا به اصل ضرب می‌شود $2!\times4! = 48$.

Scanned with CamScanner

مثال ۱۵: از یک گروه مرکب از ۵ پزشک و ۳ پرستار، چند کمیته ی ۳ نفره می توان تشکیل داد؟

حل: این گروه جمعاً ۸ نفر عضو دارد، بنابراین تعداد کمیته های ۳ نفره برابر است با $\binom{۸}{۳} = \frac{۸!}{۳!۵!} = ۵۶$.

مثال ۱۶: در مثال بالا (مثال ۱۵)، اگر بخواهیم کمیته های ۳ نفره، مرکب از ۲ پزشک و ۱ پرستار باشند، چند کمیته ۳ نفره می توان انتخاب نمود؟

حل: طبق اصل ضرب داریم:

$$\binom{۵}{۲} \times \binom{۳}{۱} = \frac{۵!}{۲!۳!} \times \frac{۳!}{۱!۲!} = ۱۰ \times ۳ = ۳۰$$

مثال ۱۷: ۴ نفر از ۷ کارگر یک کارگاه مرد هستند. به چند طریق می توان ۳ نفر از بین کارگران انتخاب نمود به طوری که حداکثر یک نفر از آن ها زن باشد؟

حل: حداکثر یک زن یعنی یا یک زن یا هیچ زن. در این سوال هم از اصل ضرب و هم از اصل جمع استفاده شده است.

$$\binom{۴}{۳}\binom{۳}{۰} + \binom{۴}{۲}\binom{۳}{۱} = \frac{۴!}{۳!۱!} \times ۱ + \frac{۴!}{۲!۲!} \times \frac{۳!}{۲!۱!} = ۴ + (۶ \times ۳) = ۲۲$$

مثال ۱۸: جعبه ای حاوی ۴ توپ قرمز، ۳ توپ سبز و ۹ توپ آبی است. به چند طریق می توان ۶ توپ را انتخاب کرد، به طوری که

الف: محدودیتی وجود نداشته باشد.

ب: ۳ توپ قرمز، ۱ توپ سبز و ۲ توپ آبی وجود داشته باشد.

حل: الف:

$$\binom{۱۶}{۶} = \frac{۱۶!}{۶!۱۰!}$$

ب:

$$\binom{۴}{۳}\binom{۳}{۱}\binom{۹}{۲} = ۴ \times ۳ \times \frac{۹!}{۲!(۹-۲)!} = ۴۳۲$$

مثال ۱۹: از ۱۰ دستگاه تلوزیون موجود در یک فروشگاه، ۳ تلوزیون نقص فنی دارند. به چند طریق صاحب یک هتل می تواند ۴ دستگاه از این تلوزیون ها را خریداری کند، به طوری که حداقل ۲ تلوزیون نقص فنی داشته باشد؟

حل: حداقل ۲ یعنی ۲ یا ۳ تلوزیون نقص فنی داشته باشند.

$$\binom{۳}{۲}\binom{۷}{۲} + \binom{۳}{۳}\binom{۷}{۱} = \left(۳ \times \frac{۷!}{۲! \times ۵!}\right) + (۱ \times ۷) = (۳ \times ۲۱) + ۷ = \boxed{۷۰}$$

1

احتمال :

احتمال پیشامد یا A را با $P(A)$ نمایش می‌دهند و به صورت زیر تعریف می‌کنیم:

$$P(A) = \frac{n(A)}{n(S)} = \frac{\text{تعداد حالات مساعد}}{\text{تعداد حالات ممکن}}$$

- احتمال یک پیشامد غیرممکن برابر صفر می‌باشد. $P(\emptyset) = 0$
- احتمال یک پیشامد حتمی برابر یک می‌باشد. $P(S) = 1$
- $0 \le P(A) \le 1$

مثال ۱: احتمال اینکه در پرتاب یک تاس، عدد رو شده زوج باشد؟ حل:

$$P(A) = \frac{n(A)}{n(S)} = \frac{3}{6} = \frac{1}{2}$$

2

احتمال شرطی: احتمال رخ دادن پیشامد A، به شرط آنکه پیشامد B رخ داده است را احتمال شرطی می‌نامند و آن را با نماد $P(A|B)$ نمایش می‌دهند و می‌خوانند، احتمال وقوع A به شرط آنکه B قبلاً رخ داده است و مقدار آن از رابطه‌ی زیر بدست می‌آید:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

اگر جای A و B عوض شود داریم:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) \neq 0$$

مثال: دو تاس را پرتاب می‌کنیم، در صورتی که بدانیم یکی از تاس‌ها عدد ۵ را نشان داده، مطلوب است محاسبه‌ی احتمال اینکه تاس دیگر عدد ۶ را نشان دهد.

حل: پیشامد A : اینکه یکی از تاس‌ها عدد ۵ را نشان دهد.

پیشامد B : " " " ۶ " " " .

بنابراین

$$A = \{(1,5), (2,5), (3,5), (4,5), (5,5), (6,5), (5,1), (5,2), (5,3), (5,4), (5,6)\}$$

$$B = \{(1,6), (2,6), (3,6), (4,6), (5,6), (6,6), (6,1), (6,2), (6,3), (6,4), (6,5)\}$$

$$A \cap B = \{(5,6), (6,5)\}$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{11}{36}, \quad P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{2}{36}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{2/36}{11/36} = \frac{2}{11}$$

درستی

روش دیگر: $S = \{(1,5)(2,5)(3,5)(4,5)(5,5)(6,5)(5,1)(5,2)(5,3)(5,4)(5,6)\}$ فضای نمونه

و پیشامد مورد نظر برابر است با $\{(5,6)(6,5)\}$ ، بنابراین احتمال برابر $\frac{2}{11}$ است.

مثال: دو تاس را پرتاب می‌کنیم، در صورتی که برآمد مجموع اعداد ظاهر شده برابر ۶ است، مطلوب است محاسبه احتمال اینکه یکی از تاس‌ها عدد ۲ را نشان دهد.

3

حل: پیشامد A : اینکه مجموع دو عدد ظاهر شده برابر ۶ است.

پیشامد B : اینکه یکی از تاس‌ها عدد ۲ را نشان دهد.

$$A = \{(1,5)(2,4)(3,3)(4,2)(5,1)\}$$

$$B = \{(1,2)(2,2)(3,2)(4,2)(5,2)(6,2)(2,1)(2,3)(2,4)(2,5)(2,6)\}$$

$$A \cap B = \{(2,4)(4,2)\}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{2/36}{5/36} = \frac{2}{5}$$

نتیجه: اگر پیشامدهای A و B بتوانند هم‌زمان امکان‌پذیر باشند، آنگاه $P(A \cap B) = P(B|A)P(A)$

تعریف: دو پیشامد A و B را مستقل گوییم، اگر و فقط اگر $P(A\cap B) = P(A)P(B)$

* با توجه به این رابطه و فرمول احتمال شرطی نتیجه می‌گیریم که $P(B|A) = P(B)$ و $P(A|B) = P(A)$

به عبارت دیگر اگر رخ دادن پیشامد A، اثری در رخ دادن یا ندادن پیشامد B نداشته باشد، دو پیشامد را مستقل می‌نامیم.

مثال: دو تاس را پرتاب می‌کنیم. مطلوب است احتمال اینکه هر دو ۶ بیایند.

حل: فرض کنید A پیشامد ۶ آمدن تاس اول، B پیشامد ۶ آمدن تاس دوم باشد.

$$P(A\cap B) = P(A)P(B) = \frac{1}{6}\times\frac{1}{6} = \frac{1}{36}$$ (چون A و B مستقل‌اند.)

مثال: احتمال اینکه تیر علی به هدف بخورد $P(A) = 0.9$ و احتمال اینکه تیر حسن به هدف بخورد $P(B) = 0.8$ می‌باشد. اگر علی و حسن با هم به هدف تیراندازی کنند، مطلوب است احتمال اینکه لااقل یک تیر به هدف بزند.

حل: پیشامدهای A و B مستقل هستند، یعنی $P(A\cap B) = P(A)P(B) = 0.72$. حال احتمال پیشامد $A\cup B$ را حساب می‌کنیم.

$$P(A\cup B) = P(A) + P(B) - P(A\cap B) = 0.9 + 0.8 - 0.72 = \boxed{0.98}$$

فرمول احتمال کل: اگر فضای نمونه‌ای S به n پیشامد $B_1, B_2, \dots, B_n$ افراز شده باشد و اگر A پیشامدی از S باشد، آنگاه به شرط $P(B_i) \neq 0$ و $i = 1, \dots, n$،

$$\boxed{P(A) = \sum_{i=1}^{n} P(A\cap B_i) = \sum_{i=1}^{n} P(B_i)P(A|B_i)}$$

مثال: سه ظرف مشابه داریم. اولین ظرف شامل ۵ مهره سفید و ۱۱ مهره سیاه است. دومین ظرف شامل ۳ مهره سفید و ۹ مهره سیاه و سومین ظرف فقط شامل مهره‌های سفید است. با چشم بسته یکی از سه ظرف را انتخاب و از آن مهره‌ای بیرون می‌آوریم. احتمال اینکه مهره سفید باشد، چقدر است؟

حل: پیشامد استخراج مهره سفید را با A، انتخاب ظرف اول $= B_1$، انتخاب ظرف دوم $= B_2$ و انتخاب ظرف سوم $= B_3$.

بدیهی است $P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$. از طرفی $P(A|B_1) = \frac{5}{16}$ ، $P(A|B_2) = \frac{1}{4}$ ، $P(A|B_3) = 1$.

حال با توجه به فرمول احتمال کل

$$P(A) = P(B_1)\cdot P(A|B_1) + P(B_2)\cdot P(A|B_2) + P(B_3)\cdot P(A|B_3)$$
$$= \left(\frac{1}{3}\times\frac{5}{16}\right) + \left(\frac{1}{3}\times\frac{1}{4}\right) + \left(\frac{1}{3}\times 1\right) = \boxed{\frac{25}{48}}$$

قاعده بیز : می‌دانیم اگر A و B دو پیشامد با احتمال مثبت از فضای نمونه‌ای یک آزمایش تصادفی باشند، آنگاه داریم :

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \quad , \quad P(B|A)=\frac{P(A\cap B)}{P(A)} \Rightarrow P(A\cap B)=P(A)\,P(B|A)$$

از دو رابطه‌ی بالا نتیجه می‌گیریم که :

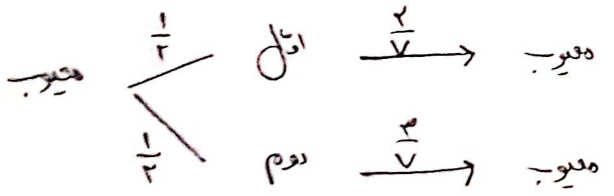
$$\boxed{P(A|B)=\frac{P(A)}{P(B)}\cdot P(B|A)}$$

این رابطه را قاعده‌ی بیز می‌نامند.

مثال : در جعبه‌ای دو ظرف وجود دارد. در ظرف اول ۵ لامپ سالم و ۲ معیوب و در ظرف دوم ۴ لامپ سالم و ۳ لامپ معیوب وجود دارد. از این جعبه لامپی به تصادف انتخاب می‌کنیم و بعد از امتحان کردن آن متوجه شدیم که لامپ معیوب است. چقدر احتمال دارد این لامپ از ظرف اول انتخاب شده باشد؟

حل : فرض کنید A احتمال معیوب بودن لامپ و B احتمال انتخاب لامپ از ظرف اول.

این سوال در واقع قاعده بیز است.



$$P(\text{ظرف اول} \mid \text{معیوب})=\frac{P(\text{ظرف اول})\,P(\text{معیوب} \mid \text{ظرف اول})}{P(\text{معیوب})}$$

$$P(B|A)=\frac{P(B)\;P(A|B)}{P(A)}=\frac{\frac{1}{2}\times\frac{2}{7}}{\left(\frac{1}{2}\times\frac{2}{7}\right)+\left(\frac{1}{2}\times\frac{3}{7}\right)}=\frac{2}{5}$$

* حالت کلی قاعده بیز به صورت زیر است :

اگر $\{H_1, H_2, \dots, H_n\}$ افرازی از S باشند، و A پیشامد دلخواهی از S باشد و $P(A)\neq 0$، آنگاه به ازای هر $j=1,2,\dots,n$ داریم :

$$P(H_j|A)=\frac{P(H_j)\,P(A|H_j)}{\sum_{i=1}^{n} P(H_i)\,P(A|H_i)}$$

متغیرهای تصادفی

تعریف متغیر تصادفی : تابعی است از فضای نمونه ای به مجموعه ای از اعداد حقیقی . یعنی $X : S \longrightarrow \mathbb{R}$.

مثال ۱ : ظرفی محتوی ۶ توپ سفید و ۵ توپ سیاه ، ۳ توپ را به تصادف و بدون جایگذاری خارج می کنیم . فرض کنید متغیر تصادفی X نشان دهنده ی تعداد توپ سفید در این آزمایش باشد . فضای نمونه آزمایش و مقادیر متغیر تصادفی X را مشخص نمایید .

حل : فضای نمونه بصورت زیر می باشد ،

$$S = \{www, wwb, wbw, bww, bbw, bwb, wbb, bbb\}$$

w ← سفید

b ← سیاه

$X(www) = ۳$

$X(wwb) = X(wbw) = X(bww) = ۲$

$X(wbb) = X(bwb) = X(bbw) = ۱$

$X(bbb) = ۰$

مقادیری که این متغیر تصادفی اختیار می کند ، اعداد ۰ ، ۱ ، ۲ و ۳ می باشد .

مثال ۲ : یک جفت تاس را پرتاب می کنیم . فرض کنید متغیر تصادفی X نشان دهنده مجموع دو عدد ظاهر شده باشد ، فضای نمونه آزمایش و مقادیر تصادفی X را مشخص نمایید .

حل : فضای نمونه آزمایش شامل ۳۶ زوج مرتب (۶×۶) می باشد .

$$S = \{(۱,۱), (۱,۲), \ldots, (۲,۱), (۲,۲), \ldots, (۶,۶)\}$$

لذا مقادیری که متغیر تصادفی X اختیار می کند ، اعداد ۲ ، ۳ ، ۴ ، ۵ ، ۶ ، ۷ ، ۸ ، ۹ ، ۱۰ ، ۱۱ و ۱۲ می تواند باشد .

توزیع برنولی: فرض کنید در یک آزمایش، پیشامد A امکان وقوع داشته باشد، به طوری که $p(A)=p$ و امکان عدم وقوع آن $p(A')=q$. واضح است که A و A' دو پیشامد متمم هستند و

$$p(A)+p(A')=p+q=1$$

چنین آزمایشی را، آزمایش برنولی می نامند.

در این مبحث، آزمایش هایی را می خواهیم مورد بررسی قرار دهیم که به یکی از دو پیشامد پیروزی یا شکست منتهی شوند. (مانند A یا A'). اگر متغیر تصادفی X را طوری تعریف کنیم که اعداد ۱ و ۰ را به ترتیب برای پیروزی و شکست قبول کند، یعنی وقتی نتیجه ی آزمایش پیروزی است $X=1$ و وقتی نتیجه ی آزمایش شکست باشد $X=0$. به عبارت دیگر اگر متغیر تصادفی X نشان دهنده ی تعداد پیروزی در یک آزمایش باشد، آنگاه تابع احتمال X عبارت است از:

$$f(1)=p(X=1)=p \quad , \quad f(0)=p(X=0)=1-p$$

که در آن p، $0\le p\le 1$ احتمال پیروزی است.

$$f(x)=p^{x}(1-p)^{1-x} \quad , \quad x=0,1$$

★ میانگین و واریانس متغیر تصادفی برنولی عبارت است از:

$$\mu=p \quad , \quad \sigma^2=p(1-p)$$

تعریف امید ریاضی (میانگین): امید ریاضی یک متغیر تصادفی، مقداری است که به طور متوسط (میانگین) حاصل می شود.

$$E(X)=\sum_{i=1}^{k} x_i p_i$$

به عبارت دیگر، برای محاسبه ی $E(X)$، هر مقدار ممکن متغیر تصادفی را در احتمال آن مقدار ضرب می کنیم و این حاصل ضرب ها را با هم جمع می کنیم.

مثال ۳: توزیع احتمال متغیر تصادفی X به صورت زیر مفروض است. مطلوب است محاسبه ی $E(X)$؟

x	۰	۱	۲	۳
$p(x)$	۰٫۴	۰٫۱	۰٫۳	۰٫۲

حل:

$$E(X)=(0\times 0.4)+(1\times 0.1)+(2\times 0.3)+(3\times 0.2)=1.3$$

تعریف واریانس: واریانس X با توجه به تعریف امید ریاضی X به صورت زیر تعریف می شود؟

$$Var(X)=E(X^2)-(E(X))^2$$

که همیشه $E(X^2)\ge E^2(X)$.

توزیع دوجمله‌ای منفی: اگر آزمایش برنولی را آنقدر ادامه دهیم تا r موفقیت بدست آید و پس از رخ دادن r-امین موفقیت، آزمایش متوقف شود، چنین آزمایشی را آزمایش دوجمله‌ای منفی و متغیر تصادفی X، که نشان دهنده‌ی تعداد دفعات در آزمایش دوجمله‌ای منفی است را، متغیر تصادفی دوجمله‌ای منفی می‌نامند.

* منظور ما در این بحث، پیدا کردن احتمال r-امین موفقیت در n-امین آزمایش است.

تعریف: توزیع متغیر تصادفی دوجمله‌ای منفی X را به صورت زیر تعریف می‌کنند؛ (p احتمال موفقیت در هر آزمایش می‌باشد.)

$$\boxed{P(X=n)=\binom{n-1}{r-1}p^r q^{n-r}} \quad , n=r, r+1, \dots$$

مثال ۴: سکه‌ای طوری اریب شده است که احتمال آمدن شیر $\frac{2}{3}$ است. احتمال اینکه چهارمین شیر در هفتمین پرتاب ظاهر شود، چقدر است؟

حل: توزیع دوجمله‌ای منفی با $p=\frac{2}{3}$ ، $r=4$ و $n=7$ است. پس

$$P(X=7)=\binom{7-1}{4-1}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^{7-4}.$$

مثال ۵: از ظرفی محتوی ۴ مهره قرمز و ۳ مهره سفید، مهره‌ای به تصادف بیرون آورده و رنگ آن را مشاهده و پس به ظرف برمی‌گردانیم. مطلوب است حساب احتمال آنکه در آزمایش هشتم، برای پنجمین بار رنگ مهره قرمز باشد.

حل: توزیع دوجمله‌ای منفی با $p=\frac{4}{7}$ ، $r=5$ و $n=8$ است. پس

$$P(X=8)=\binom{8-1}{5-1}\left(\frac{4}{7}\right)^5\left(\frac{3}{7}\right)^{8-5}.$$

اگر در آزمایش $r=1$ باشد، یعنی آزمایش برنولی را آنقدر ادامه دهیم تا اولین پیروزی حاصل شود، آنگاه آزمایش را آزمایش هندسی می‌نامیم. یعنی $n=k$ است اگر $k-1$ آزمایش اول شکست و k-امین آزمایش پیروزی باشد.

تعریف: متغیر تصادفی X که نشان دهنده تعداد دفعات آزمایش هندسی است را متغیر تصادفی هندسی می‌نامند.

تعریف: توزیع متغیر تصادفی هندسی را توزیع هندسی می‌نامند و آن را به صورت زیر تعریف می‌کنند؛

$$\boxed{P(X=n)=p\,q^{n-1}} \quad , n=1, 2, \dots$$

مثال ۶: یک تاس را به طور مکرر پرتاب می‌کنیم؛

الف: احتمال اینکه اولین عدد ۵ در ششمین پرتاب ظاهر شود، چقدر است؟ حل: این توزیع هندسی با $n=6$ و $p=\frac{1}{6}$ است پس

$$P(X=6)=\frac{1}{6}\times\left(\frac{5}{6}\right)^5=0.067$$

ب: حداقل ۶ آزمایش لازم است، تا اولین عدد ۵ ظاهر شود. حل:

$$P(X\geq 6)=\sum_{n=6}^{\infty}\left(\frac{5}{6}\right)^{n-1}\left(\frac{1}{6}\right)=\left(\frac{5}{6}\right)^5=0.402$$

پ: حداکثر ۵ آزمایش لازم است تا اولین عدد ۵ ظاهر شود. حل:

$$P(X\leq 5)=\sum_{n=1}^{5}\left(\frac{5}{6}\right)^{n-1}\left(\frac{1}{6}\right)=0.598$$

①

فصل چهارم : برآورد حدود اطمینان :

تخمین فاصله ای میانگین یک جامعه : حالت اول :

$$\mu = \bar{x} \pm \left(Z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right)$$

به عبارت دیگر

$$\bar{x} - \left(Z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right) \leq \mu \leq \bar{x} + \left(Z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right)$$

یا

$$\mu \in \left(\bar{x} - \left(Z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right) , \bar{x} + \left(Z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right)\right)$$

مثال : اگر نمونه ای تصادفی به اندازه ی $n=20$ از یک جامعه ی نرمال با واریانس $\sigma^2 = 225$ دارای میانگین $\bar{x} = 64.3$ باشد ، یک فاصله ی اطمینان ٪۹۵ برای میانگین μ بسازید.

حل :

$n = 20$ ، $\bar{x} = 64.3$ ، $\sigma = 15$ ، $\alpha = 1 - \text{فاصله اطمینان} = 1 - \frac{95}{100} = 0.05$

$\frac{\alpha}{2} = 0.05 \div 2 = 0.025$

$Z_{0.025} = 1.96$

جاگذاری در فرمول

$$64.3 - \left(1.96 \times \frac{15}{\sqrt{20}}\right) \leq \mu \leq 64.3 + \left(1.96 \times \frac{15}{\sqrt{20}}\right)$$

4.47

6.57

$$\Longrightarrow 57.7 \leq \mu \leq 70.9$$

مثال : سرعت یک مونتاژ کاری بر حسب ثانیه توسط ۲۵ نفر کارگر ماهر اندازه گیری شد و بر مبنای آن میانگین نمونه برابر ۴۳٫۷ ثانیه به دست آمد. اگر سرعت انجام این نوع مونتاژ کاری توزیع نرمال $X \sim N(\mu, 4)$ (با میانگین μ و انحراف معیار ۴) باشد ، یک فاصله اطمینان ۹۵ درصد برای μ بسازید و آن را تفسیر کنید.

حل :

$$\left(43.7 - 1.96 \times \frac{4}{5} , 43.7 + 1.96 \times \frac{4}{5}\right) = (42.13 , 45.26)$$

در نتیجه فاصله $(42.13 , 45.26)$ یک فاصله اطمینان ۹۵ درصد برای میانگین جامعه است

تفسیر : از هر صد فاصله ای که برای μ بر مبنای مکانیزم نمونه گیری ، شرح ارائه شده در مسئله حساب شود ، ۹۵ تا μ را در بر دارند و در پنج مورد از صد بار تکرار μ در فاصله نخواهد بود.

نکته: حوزه‌های استنباط آماری شامل برآورد و آزمون فرض می‌باشد. در حوزه برآورد، هدف یافتن یک عدد یا یک فاصله برای صفت مورد نظر در جامعه است که به ترتیب برآورد عددی و فاصله اطمینان را فراهم می‌کنند.

در بحث آزمون فرض، هدف ارزیابی یک ادعا درباره ویژگی مورد بررسی است.

حالت دوم: ۱- انحراف معیار جامعه نامعلوم
۲- حجم نمونه کوچک

$$\mu = \bar{X} \pm t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}$$

مثال: یک سازنده‌ی رنگ می‌خواهد متوسط زمان خشک شدن رنگ جدید دیوارهای داخلی ساختمان را معین کند. اگر برای ۱۲ سطح آزمایشی با مساحت‌های برابر، میانگین زمان خشک شدن را مساوی ۶۶٫۳ دقیقه و انحراف معیار را مساوی ۸٫۴ دقیقه به دست آورده، یک فاصله اطمینان ۹۵٪ برای میانگین واقعی μ به دست آورید.

حل: جایگذاری $\bar{X} = 66.3$ ، $S = 8.4$ ، $t_{0.025, 11} = 2.201$ فاصله‌ی اطمینان ۹۵٪ برای μ به صورت

$$66.3 - \left(2.201 \times \frac{8.4}{\sqrt{12}}\right) < \mu < 66.3 + \left(2.201 \times \frac{8.4}{\sqrt{12}}\right)$$

۵٫۳۲
۲٫۴۲ ، ۲٫۴۹

$$\Rightarrow 60.98 \le \mu \le 71.62$$

مثال: نمونه‌ای تصادفی از ۹ دانه با میانگین ۵۰ و انحراف معیار ۸ داریم. یک فاصله اطمینان ۹۵٪ برای μ بسازید.

حل:

$$50 - \underbrace{\left(2.063 \times \frac{8}{\sqrt{9}}\right)}_{5.50} < \mu < 50 + \underbrace{\left(2.063 \times \frac{8}{\sqrt{9}}\right)}_{5.50}$$

$\boxed{t_{0.025, 8} = 2.0639}$

۵٫۴۸ ، ۵٫۴۸ ✓

$$\rightarrow 44.5 < \mu < 55.5$$

مثال: طول جمجمه‌های ۲۵ اسکلت فسیل شده‌ی نوعی از پرندگان که نسل آن‌ها نابود شده است، دارای میانگین ۵٫۶۸ سانتی‌متر و انحراف معیار ۰٫۲۹ سانتی‌متر است. با فرض اینکه چنین اندازه‌هایی به طور نرمال توزیع شده‌اند، یک فاصله‌ی اطمینان ۹۵٪ برای طول میانگین جمجمه‌های این نوع پرندگان پیدا کنید.

$$\mu = 5.68 \pm \underbrace{1.96 \times \frac{0.29}{\sqrt{25}}}_{0.11}$$

۰٫۰۵۸

$$5.57 \le \mu \le 5.79$$

آزمون فرضیه‌ها

<u>آزمون فرض</u>: روشی است برای بررسی درستی یا نادرستی تخمین ارائه شده برای پارامتر θ.

<u>فرض آماری</u>: فرض آماری، حکمی است درباره‌ی جامعه. قابل قبول بودن آن باید بر مبنای اطلاعات حاصل از نمونه‌گیری از جامعه بررسی شود.

<u>فرض‌های H_0 و H_1</u>: H_0 را فرض صفر و H_1 را فرض مقابل می‌نامند. هرگاه بخواهیم درستی یا نادرستی ادعایی را به وسیله اطلاعات حاصل از نمونه تعیین کنیم، آن ادعا را فرض مقابل (H_1) و نقیض آن را فرض صفر (H_0) می‌گیریم.

<u>مثال</u>: فرض کنید متوسط محصول گندم یک بذر موجود در بازار 5 تن در هکتار بوده است. یک محقق که روی اصلاح این بذر کار کرده، مدعی است، کار او موجب افزایش محصول شده است. فرض H_0 و H_1 چگونه است؟

<u>حل</u>: با توجه به اطلاعات مسئله، ادعای محقق $\mu > 5$ (μ متوسط) است و با در نظر گرفتن تعریف H_0 و H_1 داریم:

$$\begin{cases} H_0 : \mu \leq 5 \\ H_1 : \mu > 5 \end{cases}$$

<u>مثال</u>: متوسط طول عمر یک لامپ <u>1100</u> ساعت می‌باشد. مدیر کارخانه ادعا می‌کند که متوسط طول عمر لامپ‌های او بیشتر از <u>1100</u> ساعت است. فرض H_0 و H_1 مناسب را بنویسید.

<u>حل</u>:

$$\begin{cases} H_0 : \mu \leq 1100 \\ H_1 : \mu > 1100 \end{cases}$$

<u>خطاهای آماری</u>: هر آزمون 2 نوع خطای آماری دارد؛

<u>خطای نوع I</u>: (رد ناحق H_0): رد کردن H_0 وقتی H_0 درست است. به عبارت دیگر، اگر H_0 صحیح باشد، تصمیم نادرستی که ممکن است گرفته شود، قبول H_1 می‌باشد. این تصمیم نادرست را خطای نوع I می‌نامند و احتمال اتخاذ این تصمیم نادرست را با α نشان می‌دهند.

$$\underbrace{\alpha}_{\text{سطح معنی‌داری آزمون}} = \text{احتمال خطای نوع I} = P(\text{رد } H_0 \mid \text{درستی } H_0)$$

خطای نوع II: (قبول یا ناحق H_0) رد نکردن H_0 وقتی H_1 درست است. به عبارت دیگر اگر H_1 صحیح باشد، تصمیم نادرستی که ممکن است گرفته شود، قبول H_0 می‌باشد. این تصمیم نادرست را خطای نوع II می‌نامند و احتمال اتخاذ این تصمیم نادرست را با β نمایش می‌دهند.

$$\beta = \text{احتمال خطای نوع II} = P(\text{قبول } H_0 \mid \text{نادرستی } H_0)$$

مثال: فرض کنید X دارای توزیع دوجمله‌ای با پارامترهای ۴ و p می‌باشد. برای آزمودن فرض $H_0: p=\frac{1}{3}$ در برابر $H_1: p=\frac{1}{5}$ اگر $x=0$ یا $x=1$ ملاک رد کردن فرض H_0 باشد، احتمال خطای نوع اول چقدر است؟

حل:

$$\alpha = P(\text{رد } H_0 \mid \text{درستی } H_0) = P(X=0 \mid p=\tfrac{1}{3}) + P(X=1 \mid p=\tfrac{1}{3})$$

$$= \binom{4}{0}\left(\frac{2}{3}\right)^4 + \binom{4}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^3 = \left(\frac{2}{3}\right)^4 + \frac{4}{3}\left(\frac{2}{3}\right)^3 = \underbrace{\left(\frac{2}{3}\right)^4}_{\text{فاکتور}}(1+2) = 3\left(\frac{2}{3}\right)^4$$

مثال: فرض کنید $\sum_{i=1}^{9} x_i$ دارای توزیع دوجمله‌ای با پارامترهای p و $n=9$ است. اگر برای انجام آزمون $H_0: p=\frac{1}{4}$ در برابر $H_1: p=\frac{1}{2}$ ناحیه‌ی بحرانی عبارت $\sum_{i=1}^{9} x_i \geq 6$ است، احتمال خطای نوع اول را بدست آورید. (**یادآوری:** مجموعه تمام مقادیری از آماره‌ی آزمون را که به ازای آن‌ها H_0 رد می‌شود، ناحیه بحرانی (رد) H_0 می‌گویند.)

حل:

$$\alpha = P(\text{رد } H_0 \mid \text{درستی } H_0) = P\left(\sum_{i=1}^{9} x_i \geq 6 \mid p=\tfrac{1}{4}\right)$$

$$= \binom{9}{6}\left(\frac{1}{4}\right)^6\left(\frac{3}{4}\right)^3 + \binom{9}{7}\left(\frac{1}{4}\right)^7\left(\frac{3}{4}\right)^2 + \binom{9}{8}\left(\frac{1}{4}\right)^8\left(\frac{3}{4}\right) + \binom{9}{9}\left(\frac{1}{4}\right)^9 \simeq 0.01$$

مثال: متغیر تصادفی X با توزیع $\frac{e^{-\lambda}\lambda^x}{x!}$، $x=0,1,2,\dots$ مفروض است. به منظور آزمودن فرض $H_0: \lambda \leq 0.7$ در مقابل $H_1: \lambda > 0.7$ یک مشاهده از X بدست می‌آوریم. اگر X متعلق به مجموعه $\{0,1\}$ باشد، H_0 را می‌پذیریم. اگر λ واقعاً مساوی ۱ است، با چه احتمالی فرض غلط را می‌پذیریم؟

حل:

$$\beta = P(\text{قبول } H_0 \mid \lambda=1) = P(X=0 \mid \lambda=1) + P(X=1 \mid \lambda=1) = e^{-1} + e^{-1} = 2e^{-1}$$

ضریب همبستگی: معیاری است که تحت تاثیر واحدهای اندازه‌گیری x و y نباشد و به منظور تعیین شدت همبستگی بین دو متغیر استفاده می‌شود.

* برای آزمون تصمیم در باره اینکه آیا رابطه‌ی آماری معناداری بین دو متغیر وجود دارد یا خیر، از ضریب همبستگی استفاده می‌شود.

* ضریب همبستگی، میزان تغییر در متغیر y را به ازای هر واحد تغییر در متغیر x نشان می‌دهد.

* یکی از فرمول‌های محاسبه‌ی ضریب همبستگی (پیرسون) به صورت زیر می‌باشد:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}}$$

*** ویژگی‌های ضریب همبستگی:**

① $-1 \le r \le 1$

② اگر دو متغیر مستقل باشند، ضریب همبستگی برابر صفر است اما عکس آن الزاماً برقرار نیست.

③ از ضریب همبستگی برای توصیف ارتباطات غیرخطی، استفاده نمی‌شود.

④ $r=1$ همبستگی کامل و مستقیم، $0<r<1$ همبستگی ناقص و مستقیم، $r=-1$ همبستگی کامل و معکوس و $-1<r<0$ همبستگی ناقص و معکوس. $r=0$ عدم همبستگی.

*** نکاتی در مورد ضریب همبستگی:**

① وقتی از ضریب همبستگی صحبت می‌کنیم، نمی‌توان گفت که مثبت بهتر است یا منفی. به طور کلی، هر چه قدر عدد به دست آمده صرف نظر از علامت آن، بزرگتر باشد، همبستگی بیشتر است و علامت مثبت و منفی فقط جهت همبستگی دلالت دارد.

② اگر $r=0.3$ یا $r=0.9$ باشد، این فقط بدین معناست که ضریب همبستگی در دو حالت مثبت است و نمی‌توان گفت که $r=0.9$ همبستگی‌اش سه برابر $r=0.3$ است.

③ اگر همبستگی بین دو متغیر صفر است، این بدین معنا است که این دو متغیر با هم رابطه‌ی خطی ندارند. ولی ممکن است به طور غیرخطی به یکدیگر مربوط باشند.

④ اگر $y=x+1$، آنگاه ضریب همبستگی برابر 1 خواهد بود.

⑤ اگر تمام اعداد داده‌ها با هم برابر باشند، یعنی پراکندگی صفر است و ضریب همبستگی نیز صفر است.

ضریب تعیین: ضریب تعیین معلوم می‌کند که چند درصد از تغییرات متغیر y ناشی از تغییرات متغیر x است.
اگر ضریب همبستگی را به توان ۲ برسانیم، ضریب تعیین به دست خواهد آمد، که عددی بین صفر و یک است.

به عنوان مثال: اگر ضریب تعیین دو متغیر قد و وزن برابر ۰٫۶۵ به دست آید، یعنی ما می‌توانیم ۶۵ درصد از تغییرات در متغیر وزن را با تغییرات در متغیر قد بیان کنیم.

رگرسیون : در تحلیل همبستگی، هدف ما اولیه، اندازه گیری میزان همبستگی خطی بین دو متغیر است. اما اصولاً در تحلیل رگرسیون به دنبال چنین اندازه گیری نیستیم. بلکه سعی داریم، مقدار متوسط یک متغیر را براساس مقادیر ثابت متغیرهای دیگر تخمین بزنیم یا پیش بینی کنیم. هدف پیش بینی یا بیان تغییرات یک متغیر، براساس اطلاعات متغیر دیگر.

* ما باید فقط وقتی از معادله خط رگرسیون استفاده کنیم که r یک همبستگی خطی معنادار را نشان دهد.

معادله خط رگرسیون : رگرسیون را با معادله ی رگرسیون بیان می کنند که به صورت زیر می باشد.

$$\boxed{y = \alpha + \beta x}$$

متغیر وابسته (y) — متغیر مستقل (x)

که در آن، α عرض از مبدا و ضریب ثابت است و

β شیب خط رگرسیون و ضریب زاویه (ضریب رگرسیون) می باشد.

به عنوان مثال؛ رابطه ی بین قد و وزن انسان ها را در نظر بگیرید. همه می دانیم که این رابطه یک رابطه مستقیم یا همبستگی ۱۰۰٪ نمی باشد. اما می توان گفت با احتمال خیلی قوی، افراد با قد بلندتر، وزن بیشتری نیز دارند. این مثالی از رگرسیون با دو متغیر، یکی متغیر مستقل و دیگری متغیر وابسته است.

* رگرسیون حالات و انواع دیگری نیز دارد. می توان به جای یک متغیر وابسته (وزن) و یک متغیر مستقل (قد)، یک متغیر وابسته و چندین متغیر مستقل داشت.

به عنوان مثال؛ درآمد حاصل از فروش به عواملی چون هزینه ی تبلیغات و مخارج کنترل کیفیت کالا بستگی دارد.

و یا جوش آمدن آب، به عنصر دما، به فشار آب نیز بستگی دارد.

* یکی از روابطی که می توان از طریق آن، ضریب رگرسیون (β) را محاسبه نمود، به صورت زیر می باشد:

$$\beta = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

* نحوه ی پیدا کردن ضریب ثابت یا عرض از مبدا (α):

$$\bar{x} = \frac{\sum x}{n} \quad , \quad \bar{y} = \frac{\sum y}{n} \quad , \quad \boxed{\alpha = \bar{y} - \beta \bar{x}}$$

نکاتی در مورد رگرسیون:

(۱) در رگرسیون، اگر علامت β مثبت باشد، r مثبت و اگر علامت β منفی باشد، r منفی است.

(۲) خط رگرسیون، خط حداقل مجذورات است؛ خطی است که خطاهای پیش بینی را به حداقل می رساند.

مثال ۱: ضریب همبستگی و معادله خط رگرسیون را برای داده های جدول زیر محاسبه نمایید.

x_i	۲	۳	۵	۱۰
y_i	٤	۶	۱۲	۸

حل:

x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$
۲	٤	۳-	۳-	۹	۹	۹
۳	۶	۲-	۱-	۲	٤	۱
۵	۱۰	۰	۳	۰	۰	۹
۱۰	۸	۵	۱	۵	۲۵	۱
$\sum x_i = 20$	۲۸			$\sum (x_i-\bar{x})(y_i-\bar{y}) = 14$	$\sum = 38$	$\sum = 20$

$$\bar{x} = \frac{\sum x_i}{4} = \frac{20}{4} = 5 \quad , \quad \bar{y} = \frac{28}{4} = 7$$

ضریب همبستگی
$$r_{x,y} = \frac{\sum_{i=1}^{4}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}} = \frac{14}{\sqrt{38 \times 20}} = \frac{14}{\sqrt{760}} = \frac{14}{27{,}56} = \boxed{0{,}58}$$

$$\beta = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2} = \frac{14}{38} = 0{,}42$$

$$\alpha = \bar{y} - \beta\bar{x} = 7 - (0{,}42 \times 5) = 7 - 2{,}1 = 4{,}9$$

معادله خط رگرسیون
$$y = \alpha + \beta x \Longrightarrow \boxed{y = 4{,}9 + 0{,}42\, x}$$

مثال ۲: فرض کنید X هزینه ی تبلیغ در یک ماه برحسب میلیون و Y مبلغ فروش در ماه بعد برحسب میلیون تومان باشد. اطلاعات مربوط به ۶ ماه در جدول زیر موجود باشد؛ ضریب همبستگی پیرسون و معادله خط رگرسیون آن را بنویسید.

x	۲	۶	۵	۵	۸	٤
y	۲۰	۳۰	۲۸	۳۲	٤۰	۳۰

حل:

x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$
۲	۲۰	۳-	۱۰-	۳۰	۹	۱۰۰
۶	۳۰	۱	۰	۰	۱	۰
۵	۲۸	۰	۲-	۰	۰	٤
۵	۳۲	۰	۲	۰	۰	٤
۸	٤۰	۳	۱۰	۳۰	۹	۱۰۰
٤	۳۰	۱-	۰	۰	۱	۰
$\sum x_i = 30$	۱۸۰			$\sum (x_i-\bar{x})(y_i-\bar{y}) = 60$	۲۰	۲۰۸

ادامه حل مثال ۲ :

$$\bar{x} = \frac{20}{4} = 5 \quad , \quad \bar{y} = \frac{180}{4} = 30$$

$$r = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}} = \frac{60}{\sqrt{20 \times 208}} = \frac{60}{\sqrt{4160}} = \frac{60}{64{,}49} = \boxed{0{,}93}$$

ضریب همبستگی

$$\beta = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2} = \frac{60}{20} = 3$$

$$\alpha = \bar{y} - \beta\bar{x} \Longrightarrow \alpha = 30 - (3 \times 5) = 15$$

$$y = \alpha + \beta x \Longrightarrow \boxed{y = 15 + 3x}$$

معادله خط رگرسیون

مثال 3 : براساس مشاهدات $(1,4)$ ، $(2,0)$ ، $(1,-1)$ و $(-2,1)$ معادله خط برآورد رگرسیونی y روی x را پیدا کنید.
ضریب همبستگی را نیز بدست آورید.

حل :

x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$
-2	1	$-2{,}5$	0	0	$6{,}25$	0
1	-1	$,5$	-2	-1	$,25$	4
2	0	$1{,}5$	-1	$-1{,}5$	$2{,}25$	1
1	4	$,5$	3	$1{,}5$	$,25$	9
$\sum x_i = 2$	4			$\sum(x_i-\bar{x})(y_i-\bar{y}) = -1$	$\sum = 9$	14

$$\bar{x} = \frac{2}{4} = \frac{1}{2} = ,5 \quad , \quad \bar{y} = \frac{4}{4} = 1$$

$$r = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}} = \frac{-1}{\sqrt{9 \times 14}} = \frac{-1}{3\sqrt{14}} = \frac{-1}{11{,}22} = \boxed{0{,}089 \simeq 0{,}09}$$

ضریب همبستگی

$$\beta = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2} = \frac{-1}{9}$$

$$\alpha = \bar{y} - \beta\bar{x} \Longrightarrow \alpha = 1 + \frac{1}{9} \times \frac{1}{2} = \boxed{\frac{19}{18}}$$

$$y = \alpha + \beta x \Longrightarrow \boxed{y = \frac{19}{18} + \left(\frac{-1}{9}\right)x = \frac{19}{18} - \frac{1}{9}x}$$

معادله خط رگرسیون