مثال ۷-۲: اگر مقدار دبی در سه شاخهٔ فرعی رودخانه ای به ترتیب از مقادیر q_1، q_2 و q_3 بیشتر شود، آن گاه آن رودخانه سیلابی می شود. مقادیر احتمال ایجاد سیلاب در هر شاخه و حالت های ترکیبی آن ها به شرح زیر است.

$$P(Q_1 \geq q_1) = 0.3\,,\quad P(Q_2 \geq q_2) = 0.4\,,\quad P(Q_3 \geq q_3) = 0.2$$

$$P(Q_1 \geq q_1\,, Q_3 \geq q_3) = 0.15\,, P(Q_2 \geq q_2\,, Q_3 \geq q_3) = 0.2, P(Q_1 \geq q_1, Q_2 \geq q_2) = 0.1$$

$$P(Q_1 \geq q_1\,, Q_2 \geq q_2\,, Q_3 \geq q_3) = 0.05$$

مطلوب است محاسبه احتمال این که حداقل در یکی از شاخه های رودخانه سیلاب اتفاق بیافتد.

حل: مقدار $P(Q_1 \geq q_1 \cup Q_2 \geq q_2 \cup Q_3 \geq q_3)$ یعنی احتمال این که در رودخانه حالت سیلابی ایجاد شود:

$$= P(Q_1 \geq q_1) + P(Q_2 \geq q_2) + P(Q_3 \geq q_3)$$
$$-P(Q_1 \geq q_1, Q_2 \geq q_2) - P(Q_1 \geq q_1, Q_3 \geq q_3)$$
$$-P(Q_2 \geq q_2, Q_3 \geq q_3) + P(Q_1 \geq q_1\,, Q_2 \geq q_2\,, Q_3 \geq q_3) = 0.5$$

یعنی احتمال سیلابی شدن جریان در رودخانه ۵۰ درصد خواهد بود.

۷-۳ توزیع های احتمالاتی

متغیری که ناشی از یک فرآیند تصادفی است را متغیر تصادفی (Random variable) گویند. اگر X را به عنوان یک متغیر تصادفی نظیر رواناب فرض کنیم، در زمان ها و مکان های مختلف، مقدار این متغیر تصادفی تفاوت می کند. مقدار مشاهده شده (Observed value)، مقداری است که در یک واقعه در طبیعت اندازه گیری شده است.

اگر احتمال وقوع یک فرآیند تصادفی با متغیر X برابر با مقدار مشخص a شود، آن را به صورت نشان $P(X = a)$ می دهیم و به طور مشابه $P(a \leq X \leq b)$، نمایشگر احتمال وقوع X در بازهٔ $a \leq X \leq b$ است. حال اگر مقدار $P(a \leq X \leq b)$ احتمال برای کلیهٔ مقادیر a و b را بدانیم، آن گاه اشراف کاملی نسبت به احتمال وقوع برای متغیر بین مقادیر مختلف a و bخواهیم داشت، که اصطلاحاً گفته می شود که توزیع احتمالی (Probability distribution) یا به طور خلاصه توزیع متغیر تصادفی X معلوم است.

اگر x نشانگر عددی باشد که احتمال کمتر یا مساوی بودن متغیر تصادفی از آن به صورت $P(X \leq x)$ نشان داده شود، آن گاه تابع توزیع (Distribution functions) (d.f.)به صورت زیر نشان داده می شود:

$$F(x) = P(X \leq x) \qquad (۷-۸)$$

فرض کنیم a و b مقادیر عددی با فرض $b > a$ باشند. وقایع $X \leq a$ و $a \leq X \leq b$ به طور هم زمان اتفاق نمی افتد و لذا مجموع این وقایع منتج به این خواهد شد که $X \leq b$ باشد و لذا:

$$P(X \leq b) = P(X \leq a) + P(a \leq X \leq b) \qquad (۷-۹)$$

و از این رو:

$$P(a \le X \le b) = F(b) - F(a) \qquad (۱۰-۷)$$

لذا اگر تابع توزیع F(x) برای کلیهٔ مقادیر x باشد، توزیع احتمال متغیر به طور کامل مشخص خواهد بود. از آن جا که بر اساس قوانین احتمالات، مقدار احتمال نا منفی است، لذا بر اساس رابطهٔ (۷ - ۱۰) همواره $F(b) \ge F(a)$ خواهد بود و لذا تابع توزیع در هیچ موقعی یک تابع کاهشی از متغیر نخواهد بود.

برای مقادیر بی نهایت داریم:

$$F(+\infty) = \lim_{x\to\infty} F(x) = 1 \qquad (۱۱-۷)$$

$$F(-\infty) = \lim_{x\to\infty} F(x) = 0 \qquad (۱۲-۷)$$

اگر نیز در معادلهٔ (۷-۱۰) مقدار b ثابت و مقدار a به سمت b میل کند و بار دیگر عکس این عمل انجام گیرد، داریم:

$$F(a) - F(a-0) = P(X = a) \qquad (۱۳-۷)$$

$$F(a+0)-F(a)=0 \qquad (۱۴-۷)$$

معادلهٔ (۷-۱۴) نشان می دهد که تابع توزیع F(x) از سمت راست هر مقدار X پیوسته است. از طرف دیگر معادلهٔ (۷-۱۳) نشان می دهد که F(x) از سمت چپ هر نقطه ای نظیر X=a گسسته است، زیرا که $P(x = a) > ۰$.

توزیع های احتمالاتی زیادی را می توان در فرآیندهای هیدرولوژیکی بکار برد که به طور کلی به دو دستهٔ گسسته (Discrete) و پیوسته (Continuous) تقسیم می شود.

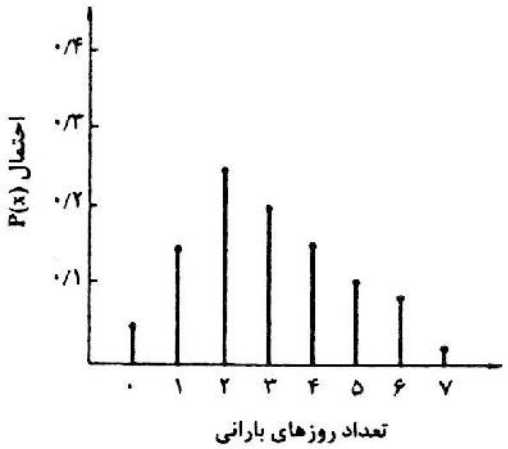
۷-۳-۱ توزیع های گسسته

متغیری را گسسته گوئیم که فقط اعداد صحیح را بتوان به آن نسبت داد. مثلاً N به عنوان تعداد روزهای بارانی در طول یک ماه، یک متغیر N گسسته نامیده می شود. زیرا نمی تواند عددی غیرصحیح یا اعشاری به آن نسبت داد.

برای مثال در نمودار ۷-۱ توزیع احتمالاتی تعداد روزهای بارانی را در طول یک ماه برای محل مشخص نشان داده شده است. این توزیع گسسته است، زیرا تعداد روزها عدد صحیح است. بر اساس اطلاعات تاریخی از تعداد روزهای بارانی در این ماه، مشخص شده که بیش از ۷ روز بارانی در این ماه اتفاق نیفتاده و احتمال بارانی بودن یک تا هفت روز برابر است با:

$P(۰) = ۰/۰۵,\ \ P(۱),\ \ P(۲) = ۰/۲۵,\ \ P(۳) = ۰/۲,\ \ P(۴) = ۰/۱۵$
$P(۵) = ۰/۱,\ \ P(۶) = ۰/۰۸,\ \ P(۷) = ۰/۰۲$

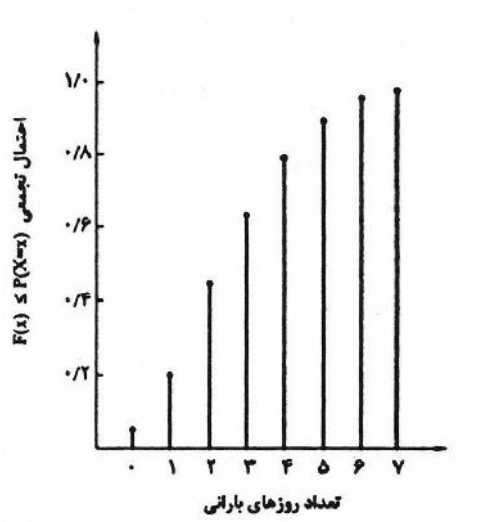
لذا مجموع این احتمالات یک خواهد بود. نحوهٔ دیگر بررسی احتمالاتی، استفاده از تابع توزیع تجمعی احتمال(Cumulative probability distribution function-CDF) به صورت $F(x) = P(X \leq x)$ است، که در جدول ۷-۱ و نمودار ۷-۲ مقادیر آن نشان داده شده است.



نمودار ۷-۱ توزیع احتمالاتی روزهای بارانی در یک ماه خاص

جدول ۷-۱ توزیع احتمالاتی و تابع توزیع تجمعی

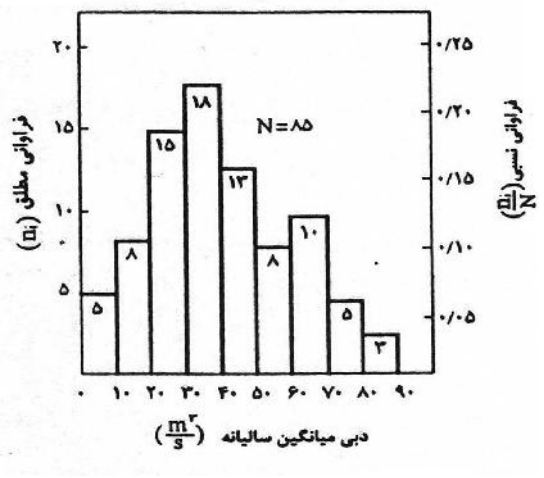
F(x)	P(x)	x
۰/۰۵	۰/۰۵	۰
۰/۲۰	۰/۱۵	۱
۰/۴۵	۰/۲۵	۲
۰/۶۵	۰/۲۰	۳
۰/۸۰	۰/۱۵	۴
۰/۹۰	۰/۱۰	۵
۰/۹۸	۰/۰۸	۶
۱/۰۰	۰/۰۲	۷



نمودار ۷-۲ تابع توزیع تجمعی احتمال روزهای بارانی

۷-۳-۲ توزیع های پیوسته

متغیری پیوسته است که بتواند هر مقدار اعشاری را به خود اختصاص دهد. به عنوان مثال، اگر دبی عبوری از مقطع معینی از رودخانه Q یک متغیر تصادفی در نظر گرفته شود، آن گاه این متغیر می تواند کلیهٔ اعداد صحیح و اعشاری را به خود اختصاص دهد، لذا یک متغیر پیوسته است. در عمل برای نشان دادن توزیع های پیوسته اگر طول کلی آن N باشد، آن را به تعدادی فواصل با بازه های مشخص تقسیم بندی می کنند. به عنوان مثال در نمودار ۷-۳ هیستوگرام ۸۵ سال آمار دبی رودخانه ای نشان داده شده است. دبی های مشاهداتی به ۹ بازه از صفر تا ۹ متر مکعب بر ثانیه تقسیم گردیده و بر این اساس تعداد n_i رخ داد، که تعداد دبی های اتفاق افتاده در هر بازه است، بر روی محور سمت چپ به عنوان فراوانی مطلق نشان داده شده است و هم چنین روی محور سمت راست مقدار فراوانی نسبی (n_i / N) نشان داده شده است. در شکل تابع CDF دبی این رودخانه رسم گردیده است.

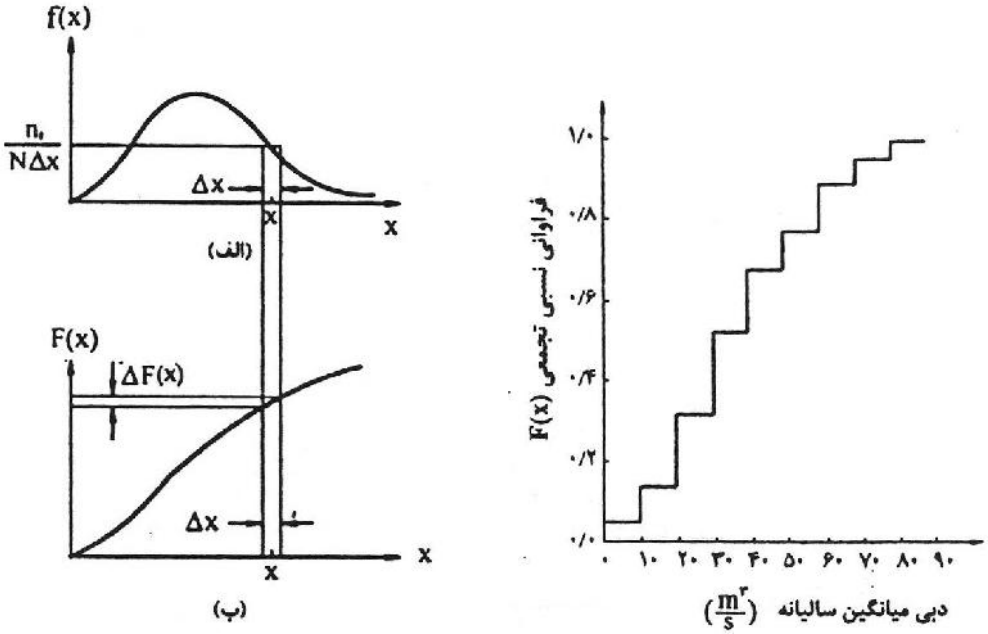


نمودار ۷-۳ توزیع فراوانی دبی میانگین سالانه

هر چه تعداد مقادیر مشاهداتی N بیشتر شود، توزیع به پیوسته نزدیک تر شده و لذا بازه های تقسیم کوچک تر می شود. لذا در نهایت نمودار های ۷-۳ و ۷-۴ تبدیل به نمودار ۷-۵ می شود، با این تفاوت که محور قائم نمودار ۷-۳ که نشانگر فراوانی نسبی بود به احتمال تبدیل شده و از این رو مساحت سطح زیر این منحنی یک می شود. بر این اساس، مقدار $\frac{n_i}{N}$ مفهوم احتمال دارد و نسبت $\frac{n_i}{N\Delta x}$ مقدار احتمال در واحد عرض است، که نشانگر چگالی میانگین احتمال است. احتمال $\frac{n_i}{N}$ در هر فاصله روی نمودار CDF نشانگر $\Delta F(x)$ یا $F(x+\frac{\Delta x}{2})-F(x-\frac{\Delta x}{2})$ است، لذا:

$$f(x)=f(x)=\lim_{\Delta x\to 0}\frac{\Delta F(x)}{\Delta x}=\frac{dF(x)}{dx} \qquad (۱۵-۷)$$

که f(x) به عنوان تابع چگالی احتمال (Probability density function) نامیده می شود. این تابع در حقیقت چگالی یا شدت احتمال در هر نقطه است و f(x)dx مقدار احتمال تفاضلی را نشان می دهد.



نمودار ۷-۴ توزیع فراوانی نسبی تجمعی دبی میانگین سالانه نمودار ۷-۵ توزیع پیوسته احتمال دبی میانگین سالانه

برای متغیرهای پیوسته همواره $f(x) \geq 0$ بوده و لذا مقدار احتمالات منفی نیست. به علاوه برای هر تابع احتمال:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \qquad (۱۶-۷)$$

که به مفهوم یک شدن مجموع احتمالات کلیهٔ رخ دادها است. احتمال این که X بین a تا b قرار گیرد برابر است با

$$P(a \leq X \leq b) = \int_a^b f(x)dx \qquad (۱۷-۷)$$

و احتمال این که x دقیقاً مقدار b را به خود بگیرد، صفر است:

$$\int_b^b f(x)dx = 0 \qquad (۱۸-۷)$$

مقدار CDF را بر اساس PDF می توان چنین تعریف نمود که:

$$P(-x \leq X \leq x) = \int_{-x}^{x} f(w)dw \qquad (۱۹-۷)$$

که در آن w یک متغیر نخودی (Dummy variable) است، لذا سطح زیر منحنی CDF بی معنی است و مقادیر روی محور قائم منحنی CDF برای تعیین مقدار احتمال به صورت زیر به کار می رود:

$$P(x_1 \leq X \leq x) = F(x_2) - F(x_1) \qquad (۲۰-۷)$$

در توزیع های گسسته نمی توان از شکل انتگرالی استفاده کرد و لذا معادلات به صورت مجموع نوشته می شود.

$$\sum_{i} f(x_i) = 1 \qquad (۲۱-۷)$$

$$P(a \le X \le b) = \sum_{x_i \ge a}^{x_i \le b} f(x_i) \qquad (۲۲-۷)$$

$$P(X \le x_j) = \sum_{i=۱}^{i=j} f(x_i) \qquad (۲۳-۷)$$

مثال ۷-۳: اگر تابع چگالی احتمال به صورت زیر باشد:

$$f(x) = \begin{cases} Ax & \text{for } 0 \le x \le 4 \\ A(8-x) & \text{for } 4 \le x \le 8 \\ 0 & \text{for other values} \end{cases}$$

مطلوب است مقدار A، $P(X \le 6)$ و تعیین تابع تجمعی احتمال (CDF).

حل: بر اساس معادلهٔ (۷-۱۶) می توان نوشت: $\int_0^{\infty} f(x)dx = 1$

$$\int_0^4 Axdx + \int_4^8 A(8-x)dx + \int_8^{\infty} 0dx = 1$$

$$A(\frac{x^2}{2})\Big|_0^4 + (\frac{-A(8-x)^2}{2})\Big|_4^8 = 1$$

لذا $A = \frac{1}{16}$ به دست می آید، حال: $P(X \le 6) = \int_0^4 Axdx + \int_4^6 A(8-x)dx = 14A = 14 \times \frac{1}{16} = \frac{7}{8}$

و در حالت کلی تابع CDF برابر است با: $F(x) = (\frac{Ax^2}{2})\Big|_0^4 + (\frac{-A(8-x)^2}{2})\Big|_4^8$

بر این اساس در جدول ۷-۲ مقادیر تابع PDF و CDF آمده است.

جدول ۷-۲ توزیع احتمالاتی و تابع توزیع تجمعی

x	f(x)	F(x)
۰	۰	۰
۱	۱/۱۶	۱/۳۲
۲	۲/۱۶	۴/۳۲
۳	۳/۱۶	۹/۳۲
۴	۴/۱۶	۸/۱۶
۵	۳/۱۶	۲۳/۳۲
۶	۲/۱۶	۲۸/۳۲
۷	۱/۱۶	۳۱/۳۲
۸	۰	۱

۷-۳-۳ تمایل مرکزی

میانگین ها

میانگین حسابی (Mean): متداول ترین تمایل مرکزی داده های مشاهداتی، مقدار میانگین حسابی (Arithmetic mean) آنهاست، که در حقیقت همان گشتاور مرتبه اول حول مبدأ است، که برای متغیرهای گسسته برابر است با:

$$E(X)=\sum_{i=1}^{n}P(x_i)x_i \qquad (۲۴-۷)$$

و برای متغیرهای پیوسته برابر است با:

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx \qquad (۲۵-۷)$$

E(X) به عنوان امید ریاضی (Expected value) مقدار متغیر تصادفی X نامیده شده و به صورت یکی از دو رابطه زیر و بر اساس داده های مشاهداتی (نمونه) تخمین زده می شود.

$$\overline{X}=\frac{1}{N}\sum_{i=1}^{N}x_i \qquad (۲۶-۷)$$

$$\overline{X}=\sum_{i=1}^{K}x_i f(x_i) \qquad (۲۷-۷)$$

که رابطهٔ (۷-۲۶) برای داده های با فراوانی برابر و رابطهٔ (۷-۲۷) برای داده های با فراوانی نسبی f(x) به کار می رود. هر چند میانگین حسابی به سادگی برآورد و از کلیهٔ داده های مشاهداتی استفاده شده و نیز نیازی به مرتب کردن داده ها ندارد، ولی دارای این نقایص نیز هست که اولاً به طور مستقیم قابل بررسی و اندازه گیری نیست و ثانیاً عدم اندازه گیری یا عدم دقت در اندازه گیری داده ها بر میزان دقت تمایل مرکزی داده ها تأثیر زیادی دارد.

مثال ۷-۴: بر اساس اندازه گیری های انجام شده در یک ایستگاه باران سنجی، مقدار بارش در طی روزهای مختلف مطابق جدول ۷-۳ ثبت گردیده است. مطلوب است محاسبهٔ میانگین بارش در این ایستگاه.

جدول ۷-۳ داده های ثبت شدهٔ بارش

ارتفاع بارش (mm)	۵	۸	۶	۲
تعداد روز	۳	۲	۴	۱

حل: از آن جا که بارش با فراوانی های نسبی متفاوت اتفاق افتاده، لذا بر اساس رابطهٔ (۷-۲۷) داریم:

$$\overline{X}=\sum_{i=1}^{k}x_i f(x_i) \quad , \quad f(x_i)=\frac{n_i}{N} \quad , \quad N=10$$

$$\overline{X}=\frac{3\times5}{10}+\frac{8\times2}{10}+\frac{6\times4}{10}+\frac{2\times1}{10}=5.7\ (mm)$$

میانگین وزنی: در محاسبهٔ میانگین داده ها اگر ارزش آن ها یکسان نباشد، عاملی به عنوان وزن به هر داده تعلق می گیرد که نشانگر اهمیت آن است. یعنی میانگین وزنی (weighted mean) برای داده های x_i با وزن w_i برابر است با:

$$\overline{X}=\frac{\sum_{i=1}^{N}w_i x_i}{\sum_{i=1}^{N}w_i} \qquad (۲۸-۷)$$

مثال ۷-۵: مطلوب است محاسبهٔ میانگین وزنی بارش در حوضهٔ آبریزی به مساحت ۳۱۰۰ کیلومتر مربع، که در آن ۶ باران سنج نصب شده و در طی یک رخ داد بارش مقادیر ارتفاع بارش را مطابق جدول ۷-۴ ثبت نموده است. مساحت تحت پوشش هر باران سنج نیز در این جدول درج شده است.

جدول ۷-۴ مقدار بارش و مساحت تحت پوشش باران سنج های مختلف

شماره باران سنج	ارتفاع بارش (mm)	مساحت تحت پوشش باران سنج (km^3)	$w_i x_i$
(۱)	(۲)	(۳)	(۴)
۱	۳۰	۳۲۳/۵	۹۷۰۵/۰
۲	۲۳	۳۱۵/۵	۷۲۵۶/۵
۳	۸	۸۶۱/۰	۶۸۸۸/۰
۴	۱۹/۲	۷۱۰/۰	۱۳۶۳۲/۰
۵	۰	۳۱۰/۰	۰
۶	۰	۵۸۰	۰
$\sum = 3100.0$			$\sum = 37481.5$

حل: بر اساس رابطهٔ (۷-۲۸) و مقادیر محاسبه شده در ستون (۴) جدول ۷-۴ داریم:

$$\overline{X} = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} W_i} = \frac{3748.5}{3100} = 12.09\ (mm)$$

میانگین هندسی: میانگین هندسی (Geometric mean) N داده به صورت زیر تعریف می شود:

$$\overline{X}_g = (x_1 x_2 x_3 ... x_N)^{\frac{1}{N}} \qquad (۲۹-۷)$$

میانگین هندسی بویژه برای محاسبهٔ میانگین متغیرهایی که به صورت توانی در معادلات وارد می شود کاربرد دارد.

میانگین هارمونیک: میانگین هارمونیک (Harmonic mean) داده به صورت زیر تعریف می شود:

$$\overline{X}_h = \frac{N}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+...+\frac{1}{x_N}} = \frac{N}{\sum_{i=1}^{N}\frac{1}{x_i}} \qquad (۳۰-۷)$$

هم چنین میانگین وزنی هارمونیک برابر است با:

$$\overline{X}_h = \frac{N}{\frac{w_1}{x_1}+\frac{w_2}{x_2}+\frac{w_3}{x_3}+...+\frac{w_N}{x_N}} = \frac{N}{\sum_{i=1}^{N}\frac{w_i}{x_i}} \qquad (۳۱-۷)$$

بایستی توجه نمود که میانگین حسابی یک سری دادهٔ مشخص، همواره از میانگین هندسی آن بزرگ تر و نیز میانگین هندسی آن ها از میانگین هارمونیک آن ها بزرگتر است.

<u>میانه Median</u>

اگر داده های مشاهداتی به صورت صعودی و یا نزولی مرتب شود و تعداد آن ها فرد باشد، میانه، عدد وسطی است و اگر تعداد داده ها زوج باشد، میانه در حقیقت میانگین دو عدد وسطی خواهد بود. از آن جا که مقادیر بیشتر از میانه و مقادیر

کم تر از میانه دارای احتمال برابر و معادل ۰/۵ هستند. لذا مقدار میانه در توزیع های آماری به عنوان یک تمایل مرکزی هم است. برای داده های دسته بندی شده ای که دارای فراوانی وقوع هستند. مقدار میانه برابر است با

$$\text{میانه} = L_1 + C\left[\frac{\frac{N}{2} - (\sum f)_1}{f_{medain}}\right] \qquad (۳۲-۷)$$

که در آن، L_1 = حد پائینی دستهٔ میانه، N= مجموع فراوانی ها، $(\sum f)_1$ = مجموع فراوانی ها در همهٔ دسته های کوچکتر از دستهٔ میانه، fmedian = فراوانی دستهٔ میانه، C= اندازهٔ دستهٔ میانه

مثال ۷-۶: مطلوب است برآورد مقدار میانه داده های دسته بندی شده بارش با فراوانی هر دسته (جدول ۷-۵).

جدول ۷-۵ بارش های دسته بندی شده و فراوانی آنها

دسته بارش (mm)	فراوانی (day)	فراوانی تجمعی (day)
۱۰۰-۱۱۰	۱۰	۱۰
۱۱۰-۱۲۰	۱۶	۲۶
۱۲۰-۱۳۰	۹	۳۵
۱۳۰-۱۴۰	۱۰	۴۵
۱۴۰-۱۵۰	۲۰	۶۵
۱۵۰-۱۶۰	۱۵	۸۰
۱۶۰-۱۷۰	۲۰	۱۰۰

حل: اگر داده ها را به صورت پیوسته در نظر بگیریم، لذا میانه در وسط فراوانی مجموع $(\frac{۱۰۰}{۲} = ۵۰)$ قرار گرفته و نیمی از داده ها در بالای آن و نیم دیگر در پائین آن قرار می گیرند. از طرفی مجموع فراوانی های چهار دستهٔ اول $(۱۰+۱۶+۹+۱۰ = ۴۵)$ است. لذا برای رسیدن به فراوانی ۵۰ نیاز به ۵ فراوانی از دستهٔ پنجم است و از آن جا که دستهٔ پنجم ۱۵۰-۱۴۰ است، لذا میانه در این دسته قرار می گیرد. با استفاده از رابطهٔ (۷-۳۲) داریم.

$$L_1 = 140\,(mm)\ ,\ N=100\ ,\ (\sum f)_1 = 45\ ,\ f_{median} = 20\ ,\ C = 10$$

$$\text{میانه} = 140 + 10\left[\frac{100/2 - 45}{20}\right] = 142.5\,(mm)$$

مثال ۷-۷: مطلوبست محاسبه میزان بارش مطابق دسته بندی و فراوانی وقوع مطابق جدول ۷-۶.

حل: بر اساس رابطه (۷-۳۲) داریم:

$$N = 40\ ,\ C = 8\ ,\ f_{median} = 12\ ,\ (\sum f)_1 = 17$$

$$\text{میانه} = 145 + 8\left[\frac{40/2 - 17}{12}\right] = 147.25\,(mm)$$