آمار و احتمالات در هیدرولوژی

۱۷-۱ مقدمه

عملکرد هر پروژه آبی بستگی به پیش‌بینی وقایع هیدرولوژیکی در آینده دارد. اطلاعات و داده‌هائی که در گذشته کسب و ثبت شده‌اند به ما کمک خواهد کرد تا تعدادی پارامتر آماری را بدست آورده و سپس از روی این پارامترها اتفاقاتی راکه ممکن است در آینده رخ دهد پیش‌بینی کنیم. آنچه در گذشته اندازه‌گیری و ثبت شده است بنام داده‌های تاریخی (historical data) یا متغیرهای تاریخی (historical variables) معروفند که نتایج پدیده‌های طبیعی و پیچیده هیدرولوژیکی بوده‌اند. لذا فرایندهای هیدرولوژیکی که در حال حاضر یا در آینده رخ خواهند داد نیز همین پدیده‌ها هستند که با زمان تغییر کرده یا تغییر خواهند کرد. برای ساده شدن مطلب رابطه رواناب (R) و بارندگی (P) را بیاد آورید که گفته شد بین مقدار بارندگی و روانابی که از آن ایجاد خواهد شد یک رابطه خطی وجود دارد و بر اساس اطلاعات گذشته این رابطه بصورت $R = a + bP$ می‌باشد که در آن دو پارامتر a و b دخالت داده شده‌اند. داده‌های تاریخی به ما کمک می‌کند که پارامترهای مذکور را طوری تعیین کنیم که با هر بارندگی در آینده بتوانیم رواناب

حاصل از آن را با دقت بیشتر تخمین بزنیم. گرچه این رابطه خود یک مدل ساده هیدرولوژیکی است، با این وجود مدل کردن دقیق سیستم‌های هیدرولوژیکی بسیار مشکل است.

در هیدرولوژی استفاده از مدل‌های فیزیکی برای پیش‌بینی وقایع آینده مقدور نبوده و کمتر مورد استقبال قرار گرفته‌اند. ما نمی‌توانیم در آزمایشگاه یک حوضه آبریز کوچک را بسازیم و وقایع هیدرولوژیکی را برای حوضه‌های آبریز طبیعی از روی آن پیش‌بینی کنیم. حتی حوضه‌های آزمایشی بزرگ و معرف هم آنطور که باید و شاید نتوانسته‌اند مفید واقع شوند. حال آن که استفاده از مدل‌های فیزیکی در هیدرولیک و سدسازی یا شبکه‌های توزیع آب و امثال آن بسیار مفید و کاربردی می‌باشند. به این دلیل در هیدرولوژی غالباً بجای مدل‌های فیزیکی از **مدل‌های مجرد** (abstract) که سیستم را بر اساس ترم‌ها و مفاهیم ریاضی تشریح می‌کند استفاده می‌شود. در این گونه مدل‌ها ورودی‌ها و خروجی‌های سیستم هیدرولوژی با یک سری معادلات ریاضی توصیف می‌شوند که برای استفاده از آن‌ها نیاز به اعداد و ارقام و پارامترهای آماری داریم. بعنوان مثال، اگر بارندگی در یک روز مشخص موردنظر بوده وبخواهیم آن را برای یک طرح یا پروژه آبی پیش‌بینی کنیم مقدار آن می‌تواند تابعی از زمان یا مکان بوده و یا این که کاملاً تصادفی باشد، لذا مدل‌های مجرد هیدرولوژیکی بسته به رویکرد آن‌ها متفاوت می‌باشند. این مدل‌ها اصولاً بر سه نوعند که عبارتند از:

- مدل‌های قطعی (deterministic models)
- مدل‌های تصادفی (stochastic models)
- مدل‌های احتمالاتی (probabilistic models)

مدل‌های قطعی سه خصوصیت عمده دارند:

(۱) - در این مدل‌ها دنباله وقوع داده‌ها (sequencing of occurrence of variables) مورد تجزیه و تحلیل قرار می‌گیرد.

(۲) - در این مدل‌ها از یک سری قواعد و قوانین حتمی (certain) استفاده شده و کاری به قوانین احتمالاتی نخواهد داشت.

(۳) - در مدل‌های قطعی شانس وقوع متغیرهای هیدرولوژی مدنظر قرار نمی‌گیرد.

حال آن که در مدل‌های احتمالاتی:

(۱) کاری به سری زمانی و دنباله وقوع داده‌ها ندارد.

(۲) فرض می‌شود که وقوع داده‌ها مستقل از زمان رخ داده‌اند.

(۳) شانس وقوع متغیرها بر اساس توابع توزیع احتمالاتی تعیین می‌شود.

(۴) فرض می‌شود که داده‌ها کاملاً تصادفی اتفاق افتاده‌اند.

مدل‌های تصادفی تقریباً حالتی بین مدل‌های قطعی و احتمالاتی می‌باشند. یعنی:

(۱) در این مدل‌ها دنباله وقوع داده‌ها هم مدنظر قرار می‌گیرد.

(۲) توزیع احتمالاتی داده‌ها ممکن است به زمان بستگی داشته یانداشته باشد.

(۳) داده‌ها ممکن است تصادفی باشند یا نباشند.

(۴) هر یک از داده‌ها ممکن است به سایر داده‌ها بستگی داشته باشد یا نداشته باشد.

از آنجائی که ما در هیدرولوژی به گمانه‌زنی و پیش‌بینی وقایع سروکار داریم، لذا در عمل یا از مدل‌های قطعی استفاده می‌شود و یا از مدل‌های تصادفی. با مدل‌های قطعی فقط می‌توان **گمانه‌زنی** (forecast) کرد حال آن که با مدل‌های تصادفی می‌توان **پیشگوئی** (Prediction) نمود. مثلاً با استفاده از مدل‌های قطعی می‌توان با دقت گمانه‌زنی کرد که مقدار تبخیر از سطح آزاد آب در یک دریاچه در یک روز مشخص حدوداً چقدر خواهد بود. اما این که بگوئیم مقدار باران در یک روز بخصوص یا در سال آینده چقدر خواهد بود یک مساله پیشگوئی است که کاملاً تصادفی است.

بنابراین در بررسی‌های هیدرولوژیکی، داده‌های آب - هواشناسی را که در گذشته اتفاق افتاده و بعضاً آن‌ها را اندازه‌گیری و ثبت کرده‌ایم گرفته و برای رسیدن به هدف آن‌ها را به لحاظ آماری تجزیه و تحلیل نموده و در مدل‌های مجرد از آن ها استفاده می‌کنیم. اما داده‌های خام نمی‌توانند مستقیماً مورد استفاده قرار گیرند بلکه لازم است درابتدا آن‌ها را بازسازی کرده و پس از پردازش پارامترهای خاصی را که مورد نیاز مدل‌ها می‌باشد از آن‌ها استخراج کنیم. برای پردازش داده‌های هیدرولوژیکی از یک سری قواعد و قوانین کلی آمار و احتمالات استفاده می‌کنیم که در این فصل به اختصار در مورد آن‌ها بحث خواهد شد.

داده‌های هیدرولوژی که اصطلاحاً به آنها ***متغیرهای هیدرولوژیکی*** گفته می‌شود اعدادی هستند که دارای دو ویژگی عمده می‌باشند: یکی از این ویژگی‌ها ***مقدار*** (magintude) و دیگری **بعد** (dimension) است. *مقدار*، نشان دهنده اندازهٔ کمّی آن عدد است. مثلاً واضح است که ۱۵ بزرگتر از ۱۰ می‌باشد امّا سوال این است که ۱۵ چی؟ آنچه به این سوال پاسخ می‌دهد ابعاد یا دیمانسیون آن کمیّت است و مسلم است که مثلاً ۱۵ روز با ۱۵ کیلومتر کاملاً متفاوت می‌باشد.

تمام پدیده‌های هیدرولوژی و هواشناسی را می‌توان با ابعاد اصلی طول (L)، جرم (M) و زمان (T) توصیف کرد. مثلاً معادله ابعادی دبی L^3T^{-1} و یا دیمانسیون سرعت LT^{-1} است. معادله‌هائی که اساس فیزیکی دارند و مشتمل بر پارامترهای مرتبط با یکدیگر می‌باشند از نظر ابعادی متوازن هستند. مثلاً دو طرف معادله بین دبی با سرعت و سطح مقطع $Q= AV$ به لحاظ ابعادی متوازن و برابر $L^3\ T^{-1}$ است. اما معادله‌های تجربی که بسیاری از فرمول‌های هیدرولوژیکی از این دسته‌اند ممکن است اساس فیزیکی نداشته و لذا طرفین معادله‌های تجربی به لحاظ ابعادی همخوانی نداشته باشند. البته باید گفت که این معادلات گرچه اساس فیزیکی ندارند ولی کاربرد آنها برای حل مسائل روزمره هیدرولوژیکی بسیار مفید می‌باشد. مثلاً ممکن است در یک منطقه بر اساس تحلیل‌های منطقه‌ای بین حداکثر دبی لحظه‌ای سیل (Q بر حسب

متر مکعب در ثانیه) و مساحت حوضه (A بر حسب کیلومتر مربع) رابطه‌ای مانند $Q = 3.6A^{0.63}$ برقرار باشد. هر چند این معادله به لحاظ فیزیکی دو طرف آن از نظر ابعادی متوازن نیست، اما از نظر کاربردی ممکن است مفید واقع گردد.

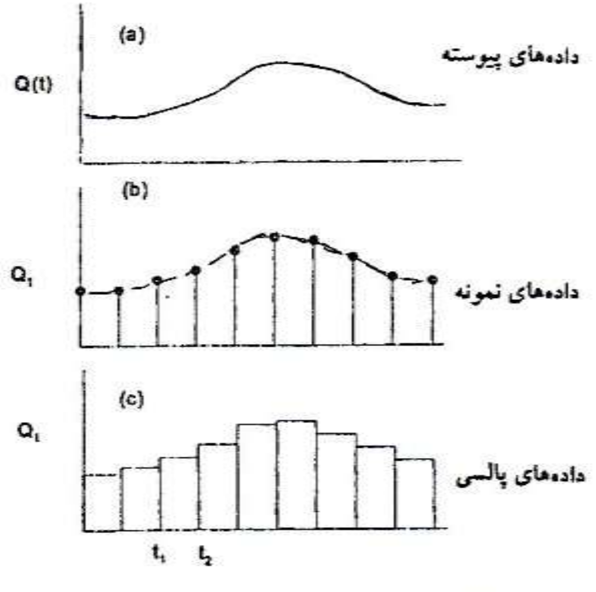
۱۷-۲ داده‌های هیدرولوژیکی

داده‌های هیدرولوژیکی مشخصه هائی هستند که از عوامل محیطی برداشت می‌شوند. داده‌ها اساساً دو گونه‌اند: ***داده‌های پیوسته*** (continuous) و ***داده‌های گسسته*** (discrete). پیوسته بودن داده‌ها به این معنی است که می‌توان بی نهایت رقم را به آن اختصاص داد. مثلاً داده‌های بارندگی پیوسته است زیرا بارش می‌تواند از نظر مقدار بی نهایت عدد داشته باشد. بعنوان مثال اگر یک روز مقدار بارندگی ۱۲ میلی متر گزارش شده است در واقع ۱۲ بصورت تصادفی رخ داده است و این مقدار می‌توانست بعنوان مثال ۱۲/۵، ۱۳، ۱۵، ۱۵/۲ و یا بی نهایت عدد دیگر باشد. یعنی پس از عدد ۱۲ تا عدد مثلاً ۱۳ بی نهایت عدد برای بارندگی می‌توانست وجود داشته باشد. ولی پاره‌ای از داده‌ها هیدرولوژیکی گسسته بوده و برای آنها فقط یک مقدار معین می‌تواند وجود داشته باشد. بعنوان مثال تعداد روزهای یخبندان در سال فقط یک عدد مشخص است و اگر این پارامتر برای یک سال مثلاً ۵۴ گزارش شده است این عدد ثابت بوده و نمی‌توانست ۵۴/۵ باشد. بعبارت دیگر بعد از ۵۴ تعداد روزهای یخبندان فقط می‌توانست ۵۵ باشد. علاوه بر این تعداد روزهای یخبندان در سال فقط بین صفر تا ۳۶۵ تغییر می‌کند. یعنی مرزهای دو طرف این داده‌ها ثابت است. تعداد انشعابات رودخانه در یک حوضه، تعداد درختان موجود در یک منطقه و یا تعداد روزهای بارانی در سال از جمله داده‌های گسسته به شمار می‌روند. بعبارت دیگر در متغیرهای گسسته بین اعداد یک گسستگی وجود دارد ولی در متغیرهای پیوسته این گسستگی وجود ندارد. زمان نیز یک متغیر پیوسته است. این وضعیت را از حرکت عقربه ثانیه شمار ساعت که بطور پیوسته در حرکت بوده و هر لحظه در یک موقعیت خاص قرار می‌گیرد مشاهده می‌کنید. طول نیز متغیر پیوسته است زیرا فاصله بین هر دو نقطه را می‌توان به بی نهایت فاصله‌های کوچک تقسیم کرد. از این نظر تجزیه و تحلیل آماری داده‌های پیوسته با داده گسسته کاملاً متفاوت است. البته در عمل متغیرهای پیوسته را چه در هنگام اندازه‌گیری و چه در هنگام تجزیه و تحلیل می‌توان بصورت داده‌های گسسته درآورد. مثلاً اگر بجای ساعت‌های معمولی از ساعت دیجیتال استفاده شود زمان بصورت اعداد منقطع روی صفحه آن نشان داده می‌شود که در واقع پیوستگی زمان را ازبین برده‌ایم.

داده‌های پیوسته بسته به روش اندازه‌گیری به دو صورت تبدیل به داده‌های گسسته می‌شوند که نتیجه آن یا ***داده‌های نمونه‌ای*** (sample data) و یا ***داده‌های پالسی*** (pulse data) است. داده‌های

نمونه‌ای نتیجه اندازه‌گیری یک پدیده در یک زمان مشخص است. مانند اندازه‌گیری دبی، درجه حرارت و یا سطح آب که مقادیر آنها در هر لحظه قابل اندازه‌گیری است. جالب آن است که ما در هیدرولوژی با داشتن تعداد معدودی داده‌های نمونه‌ای از بی‌نهایت داده‌هائی کـه مـی‌توانست وجود داشته باشد می‌خواهیم یک واقعیت هیدرولوژیکی را پیش‌بینی کنیم و این کار را بسیار مشکل و نامطمئن می‌کند.

از طرف دیگر داده‌های پالسی (تپ) عبارت است از مقدار تجمعی یک پدیده در یک فاصله زمانی مشخص (مثلاً از t_1 تا t_2).بعنوان مثال بارندگی و تبخیر و یا حجم جریان یک رودخانه از داده هائی هستند که می‌توانند بصورت پالس اندازه‌گیری شوند. شکل ۱۷-۱ نشان می‌دهد که چگونه داده‌های پیوسته دبی نسبت به زمان را می‌توان بصورت نمونه‌ای یا پالسی درآورد.



شکل ۱۷-۱ تبدیل داده‌های پیوسته به (الف) داده‌های نمونه‌ای و (ب) داده‌های پالسی.

از دیگر خصوصیات داده‌های هیدرولوژیکی سطح اندازه‌گیری است. داده‌های هیدرولوژی معمولاً بصورت پله‌ای یا زینه‌ای (interval) هستند که از صفر شروع می‌شوند. ولی صفر ممکن است همیشه به معنی صفر واقعی نباشد. مثلاً در اندازه‌گیری دمای هوا با سیستم سلسیوس دمای ۲۰ درجه دقیقاً یک درجه از ۲۱ کمتر و یک درجه از ۱۹ بیشتر می‌باشد ولی دمای ۲۰ درجه به این معنی نیست که در مقایسه با دمای ۱۰ درجه هوا دو برابـر داغ‌تـر است زیـرا صـفر درجـه سلسیوس یک صفر واقعی که در آن انرژی ماده به صفر برسد نیست. صفر درجه سلسیوس معنی فیزیکی نداشته بلکه فقط یک عدد قراردادی است. با این وجود برخی داده‌های هیدرولوژی از

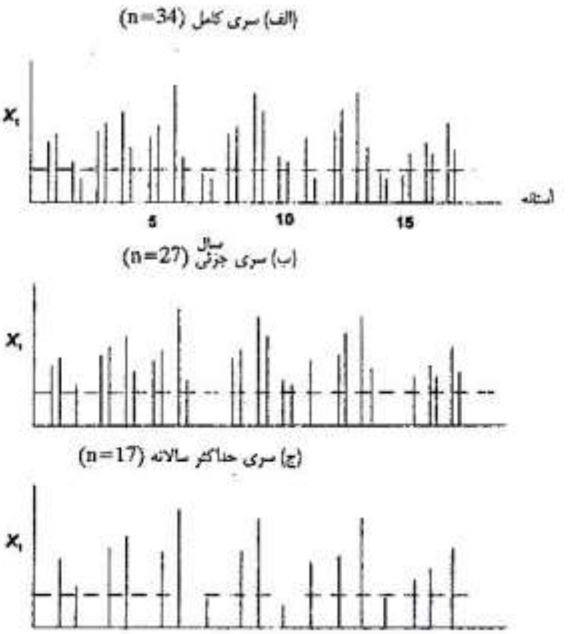
نوع نسبی (ratio) بوده و صفر آن‌ها دارای معنای فیزیکی است. مثلاً در داده‌های شدت تابش صفر به این معنی است که هیچ گونه تابشی وجود نداشته و یا مثلاً اگر شار انرژی تابشی ۵۰۰ وات بر متر مربع باشد این مقدار دقیقاً دو برابر ۲۵۰ وات بر متر مربع است. داده‌های اندازه‌گیری دما در مقیاس کلوین نیز از نوع نسبی است. زیرا صفر کلوین دارای معنی فیزیکی می‌باشد. برای توصیف داده‌های زینه‌ای یا نسبی از نمایه‌هائی مانند میانگین، انحراف از معیار و امثال آن استفاده می‌شود که به آنها پارامترهای آماری گفته می‌شود.

داده‌های هیدرولوژی را می‌توان طوری نمونه‌گیری کرد که با آن انواع سری‌ها بوجود آید. اگر داده‌ها بصورت دنباله‌وار و پشت سر هم به همان صورتی که اتفاق افتاده‌اند ثبت شود یک سری زمانی (Time Series) در اختیار خواهد بود. فرض کنید مطابق شکل ۱۷-۲ بمدت ۱۷ سال و هر سال ۲ مرتبه یک پارامتر هیدرولوژی مثل دبی (Q) را اندازه‌گیری کرده باشیم. در این صورت جمعاً ۳۴ داده خواهیم داشت که مجموعاً یک سری کامل را در طول دوره اندازه‌گیری تشکیل می‌دهد. اگر در تجزیه و تحلیل‌ها از تمام این ۳۴ عدد استفاده کنیم $(n = 34)$ آن را ***سری کامل*** (complete duration series) گویند. اما اگر برای داده‌ها یک حد مشخص را ملاک گرفته و فقط داده هائی را که بالاتر از آن هستند در نظر بگیریم در این صورت برخی از داده‌ها ممکن است حذف شوند. مثلاً در شکل ۱۷-۲ با در نظر گرفتن آستانه مورد نظر که در شکل وسط بصورت خط چین نشان داده شده است تعداد ۷ عدد حذف و تنها ۲۷ عدد دیگر $(n = 27)$ برای تجزیه و تحلیل باقی خواهند ماند. در چنین وضعیتی سری داده‌ها را ***داده‌های جزئی*** (partial duration series) گویند. مشاهده می‌شود که در این سری‌ها ممکن است داده‌های مربوط به برخی از سال‌ها بکلی حذف شوند (مانند سال‌های ۷ و ۱۴). در صورتی که شاید بهتر باشد برای هر سال حداقل یک داده در تجزیه و تحلیل دخالت داشته باشد. برای رفع این مشکل از ***سری‌های مقادیر حدّی یا اکستریم*** (extreme value series) مانند سری‌های حداکثرها یا سری‌های حداقل‌ها استفاده می‌شود که بسته به هدف بزرگترین یا کوچکترین عدد را برای هر سال در نظر می‌گیریم. مثلاً در شکل ۱۷-۲ اگر از جفت داده‌های هر سال یکی از آن‌ها راکه بزرگتر است در نظر بگیریم در اینصورت بجای ۳۴ عدد ۱۷ عدد خواهیم داشت ولی تمام سال‌ها در تجزیه و تحلیل دخالت داده شده‌اند و عددی که برای هر سال منظور شده است در بین داده‌های آن سال ماکزیمم می‌باشد. از انواع سری‌های مقادیر نهائی در هیدرولوژی ***سری‌های حداکثر سالانه*** (annual maximum series) و ***سری‌های حداقل سالانه*** (annual minimum series) است که به لحاظ هیدرولوژیکی دارای اهمیت می‌باشد.

سری‌های کامل معمولاً برای بدست آوردن منحنی تداوم جریان رودخانه‌ها و یا مطالعات خشکسالی و بارندگی مورد استفاده قرار می‌گیرد. منحنی تداوم جریان رابطه‌ای است بین دبی رودخانه و درصدی از روزهای سال که در آن دبی رودخانه بالاتر یا پائین‌تر از یک مقدار مشخص

باشد. این منحنی‌ها در برنامه ریزی‌های تأمین آب شرب و یا تامین جریان برای اهداف هیدروالکتریک مورد استفاده قرار می‌گیرد.

از طرف دیگر سری‌های جزئی و سری‌های اکستریم برای تجزیه و تحلیل فراوانی وقایع بکار برده می‌شوند. تحلیل فراوانی وقایع فرایندی است که طی آن شانس وقوع حوادث هیدرولوژی مانند سیلابهای لحظه‌ای، خشکسالی و یا امثال آن تخمین زده می‌شود.



شکل ۱۷-۲ سری‌های کامل (الف) که در آن از تمام داده‌ها استفاده شده است، سری‌های جزئی (ب) که در آن فقط داده هائی که بالاتر از یک حد مشخص است در نظر گرفته شده‌اند و (ج) سری حداکثر سالانه که در آن فقط بزرگترین داده مربوط به هر سال در نظر گرفته شده است.

داده‌های سری زمانی ممکن است ایستا (stationary) یا غیرایستا (non-stationary) باشند. داده‌ها ایستا به مجموعه داده‌هائی گفته می‌شود که خصوصیات توزیع آن‌ها مثل میانگین و انحراف از معیار در طول سری آماری ثابت بماند. بعنوان مثال چنانچه ۵۰ سال متوالی از یک ایستگاه آمار بارندگی سالانه داشته باشیم و این ۵۰ سال را به ۵ دوره ۱۰ ساله تقسیم کنیم اگر در هر یک از دوره‌های ۱۰ ساله مقادیر میانگین و انحراف از معیار حدوداً ثابت بماند گفته می‌شود که این سری زمانی ایستا می‌باشد. باید توجه داشت که سری زمانی داده‌ها اگر مربوط به یک سال باشند (سالانه) شاید ایستا باقی بمانند ولی اگر یک ماه یا یک روز را در نظر بگیریم مسلماً به دلیل نوسانات فصلی و آب و هوائی داده‌ها ایستا نخواهند بود. یعنی هر چه مقیاس را

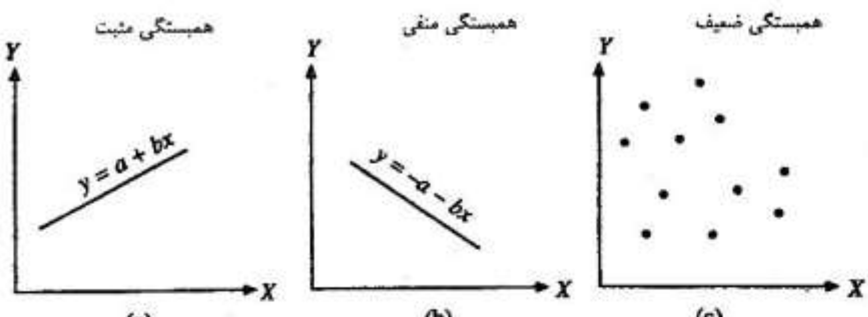
کوچکتر کنیم شانس ایستا بودن کمتر خواهد بود. ایستا بودن داده‌ها نشان از ثابت بودن شرایط هیدرولوژی در منطقه دارد.

۱۷-۳ نمودارها

در هیدرولوژی معمولاً برای نشان دادن نتایج حاصله از بررسی وضعیت داده‌ها یا متغیرهای مختلف از گراف و نمودار استفاده می‌شود. بخصوص این که امروزه نرم افزارهای گرافیکی کامپیوتری به سادگی در اختیار می‌باشند. نمودارها از این نظر حائز اهمیت هستند که علاوه بر ساده بودن بصورت عینی نیز قابل رویت است. نمودارهای x , y رابطه بین دو متغیر و نمودارهای x , y , z رابطه بین سه متغیر را مشخص می‌کنند که اولی را **نمودار دو متغیره** (bivariate) و دومی را نمودار **چند متغیره** (multivariate) گویند که در این جا از نوع سه متغیره است. نمودارهای دو متغیره را بصورت خط و نمودارهای سه متغیره را می‌توان با یک سطح نشان داد ولی روابط چند متغیره را نمی‌توان بصورت نمودار یا گراف نشان داد.

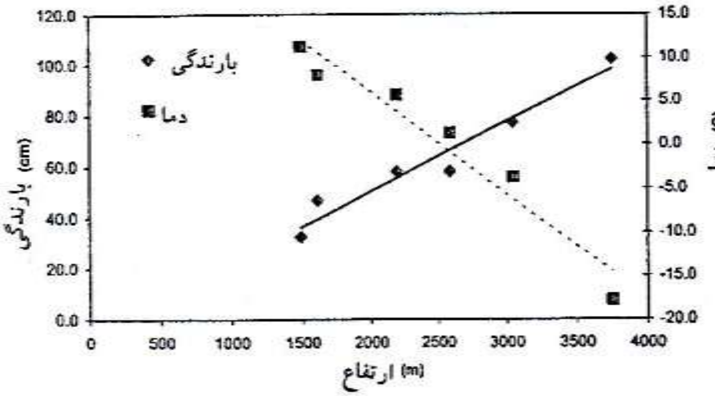
اگر یک رابطه علت و معلولی دارای دو متغیر باشد در این صورت هر تغییر در متغیر مستقل x باعث تغییراتی در متغیر وابسته y خواهد شد. شدت ارتباط بین این دو متغیر را **همبستگی** (correlation) گویند. ضریب همبستگی بین ۱+ تا ۱- تغییر می‌کند که ۱+ بمعنی همبستگی کامل مستقیم و ۱- بمعنی همبستگی کامل معکوس و ۰ بمعنی عدم همبستگی است.

در هیدرولوژی از روابط علت و معلولی زیاد استفاده می‌شود. مثلاً ممکن است بتوان دبی در پائین دست یک رودخانه را به دبی در بالا دست ارتباط داد و یا از همبستگی (correlation) بین بارندگی در یک ایستگاه با بارندگی در ایستگاه دیگر استفاده کرده و در صورت فقدان برخی داده‌های اندازه‌گیری شده در یک ایستگاه از روی این همبستگی داده‌های موردنظر تخمین زد. همبستگی بین دو پارامتر هیدرولوژیکی x , y ممکن است مطابق شکل ۱۷-۳ مثبت یا منفی بوده و یا اصولاً بین آن‌ها همبستگی وجود نداشته باشد.



شکل ۱۷-۳ همبستگی مثبت، منفی و ضعیف

در شکل ۱۷-۴ رابطه بین ارتفاع از سطح دریا با متوسط سالانه دمای هوا و بارندگی سالانه برای یک حوضه نشان داده شده است. مشاهده می‌شود که در این حوضه هر چه ارتفاع افزایش یابد مقدار بارندگی افزایش (همبستگی مستقیم) و دمای هوا کاهش (همبستگی معکوس) می‌یابد. حال در شکل ۱۷-۴ کدام متغیر مستقل و کدام متغیر وابسته است. در اینجا علت و معلول بسیار مشخص هستند زیرا این تغییر در ارتفاع است که باعث تغییر در دما یا بارندگی می‌شود پس ارتفاع متغیر مستقل و بارندگی و دما متغیر وابسته می‌باشد ولی تشخیص این دو در سایر موارد ممکن است آسان نباشد.



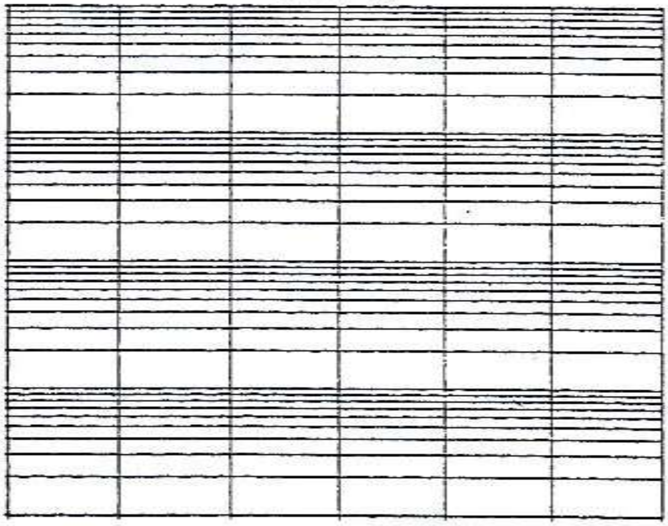
شکل ۱۷-۴ نمونه‌ای از رابطه بین بارندگی سالانه و متوسط دمای سالانه هوا با ارتفاع از سطح دریا در یک حوضه آبریز.

در شکل ۱۷-۴ محورها همگی بصورت حسابی یا ساده درجه‌بندی شده‌اند. بعبارت دیگر فاصله تقسیم‌بندی‌ها در محور x و یا محورهای y با هم برابرند. اگر دامنه تغییرات در محدوده ۱۰ تا ۱۰۰ برابر باشد امکان این که آن‌ها را روی نمودارهایی که محور آن‌ها بصورت حسابی درجه‌بندی شده باشد نشان دهیم وجود دارد ولی دامنه تغییرات برخی از داده‌های هیدرولوژی ممکن است تا یک میلیون برابر و حتی بیشتر برسد در این صورت استفاده از نمودارهای ساده امکان‌پذیر نمی‌باشد. مثلاً غلظت آلاینده‌ها در آب رودخانه‌ها از چند قسمت در بلیون (ppb) تا چند قسمت در هزار ممکن است متغیر باشد. در این صورت برای نشان دادن آن‌ها بصورت نمودار دو راه حل وجود دارد: یکی این که داده‌ها را به فرم دیگری تبدیل کنیم و دیگر آن که محورهای نمودار را به شکل دیگری تبدیل نمائیم.

اگر دامنه تغییرات داده‌ها از ۱۰۰ برابر بیشتر باشد می‌توان بجای استفاده از داده‌ها از لگاریتم آنها استفاده کرد. راه حل ساده برای چنین تبدیل هائی استفاده از کاغذ لگاریتمی است. کاغذهای لگاریتمی یا از نوع **نیمه لگاریتمی** (semi-log or arithmetic-log) هستند که در آنها یکی از محورها

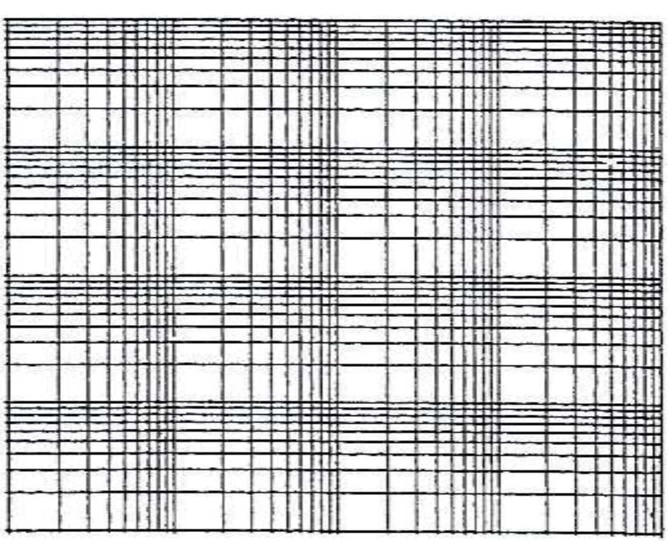
بصورت ساده و دیگری بصورت لگاریتمی درجه‌بندی شده است و یا از نوع لگاریتمی مضاعف (log-log or double log) هستند که در آن‌ها هر دو محور لگاریتمی است. شکل ۱۷-۵ کاغذ نیمه لگاریتمی و شکل ۱۷-۶ کاغذ لگاریتمی با ۴ سیکل لگاریتمی را نشان می‌دهند.

Four-cycle, semi-log



شکل ۱۷-۵ یک کاغذ گراف نیمه لگاریتمی که محور عمودی لگاریتمی است

Four-cycle, log-log



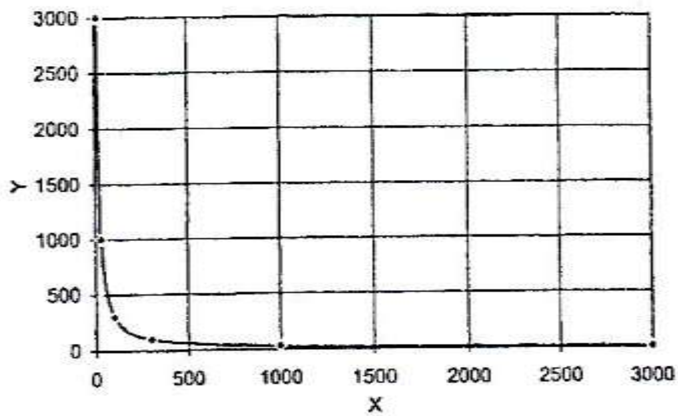
شکل ۱۷-۶ یک کاغذ گراف لگاریتمی (۴ سیکلی)

فرض کنید مقادیر متغیرهای مستقل (x) و وابسته (y) که فرضاً داده‌های هیدرولوژی می‌باشند بصورت جدول ۱۷-۱ در اختیار است. اگر بخواهیم این داده‌ها را روی محورهای مختصات نمودار حسابی رسم کنیم شکلی مطابق ۱۷-۷ بدست خواهد آمد که در آن مشاهده داده‌ها به دلیل دامنه وسیع تغییرات بسیار مشکل است.

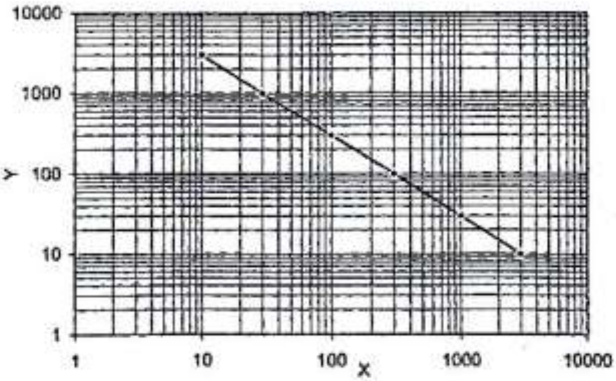
جدول ۱۷-۱ نمایش داده‌ها و تبدیل آنها به لگاریتم

X	Y	$\ln X$	$\ln Y$
10	3000	2.3026	8.0064
30	1000	3.4012	6.9078
100	300	4.6052	5.7038
300	100	5.7038	4.6052
1000	30	6.9078	3.4012
3000	10	8.0064	2.3026

یکی از راه حلها این است که از x و y لگاریتم بگیریم که در این صورت داده‌ها مطابق اعداد ستون‌های ۳ و ۴ جدول ۱۷-۱ خواهند بود. حال اگر بجای این که x و y را روی کاغذ گراف معمولی رسم کنیم. ln(x) و ln(y) را روی همین کاغذ رسم کنیم (اعداد ستون ۳ نسبت به اعداد ستون ۴) مشاهده داده‌ها و روند تغییرات آنها ساده‌تر خواهد بود. راه حل دیگر استفاده از کاغذ لگاریتمی است. در این صورت داده‌های ستون‌های ۱ و ۲ جدول ۱۷-۱ را مستقیماً روی کاغذ گراف لگاریتمی رسم می‌کنیم که شکلی مطابق ۱۷-۸ بدست خواهد آمد. در هیدرولوژی بسته به نوع داده‌ها هم از نمودارهای ساده (مثلاً در بارندگی) و هم از نمودارهای لگاریتمی (مثلاً در مورد رسوب) استفاده می‌شود.



شکل ۱۷-۷ نمودار حاصل از رسم داده‌های جدول ۱۷-۱ روی کاغذ گراف معمولی



شکل ۱۷-۸ نمودار حاصل از رسم داده‌های جدول ۱۷-۱ روی کاغذ گراف لگاریتمی

۱۷-۴ پارامترهای آماری

در این جا هدف آن نیست که آموزش آمار و احتمالات داده شود زیرا تصور می‌شود که خوانندگان با این گونه مفاهیم آشنائی دارند و در واقع آشنائی کلی با اصول آمار و احتمالات از پیش‌نیازهای هیدرولوژی کاربردی است. با این وجود لازم است به برخی روش‌ها و اصول آماری که در تجزیه و تحلیل داده‌های هیدرولوژیکی بکار گرفته می‌شود اشاره شود.

برای توصیف و محاسبه پارامترهای آماری داده‌های هیدرولوژیکی از روش موسوم به گشتاورها (method of moments) استفاده می‌شود. به این صورت که پارامترهائی مانند میانگین (mean)، واریانس (variance)، چاولگی (skewness) و کشیدگی (kurtosis) را از روی تعداد معدودی داده‌های ثبت شده که در اختیار داریم محاسبه کرده سپس تخمین‌های حاصله از نمونه‌ها را به تخمین‌های مطمئن‌تر (unbiased estimate) تبدیل کرده و برای مطالعات بعدی مورد استفاده قرار می‌دهیم. برای اطلاع از چگونگی توزیع داده‌ها باید حداقل چهار نوع مشخصه را تعیین کرد. این مشخصه‌ها عبارتند از:

(۱) پارامترهای مربوط به گرایش به مرکز (centeral tendency) مشتمل برمیانگین (mean)، میانه (median) و نما یا مد (mode).

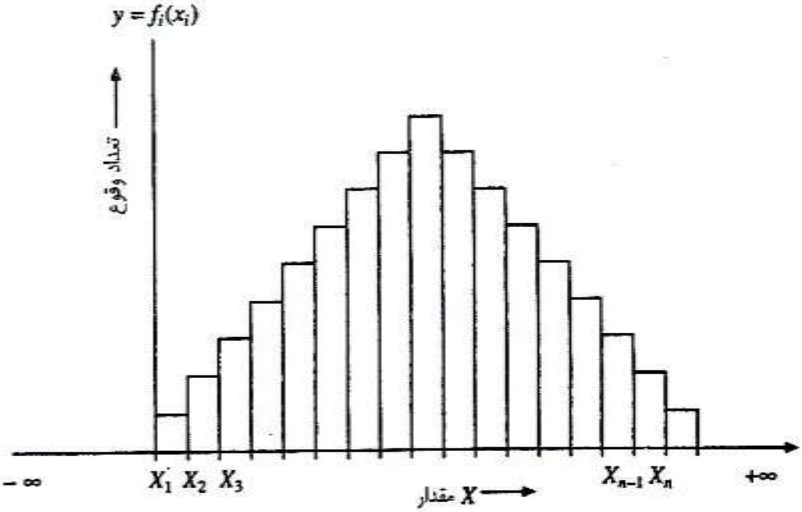
(۲) پارامترهای مربوط به پراکندگی داده‌ها مشتمل بر انحراف از معیار، ضریب تغییرات خطای معیار و دامنه داده‌ها

(۳) چاولگی

(۴) کشیدگی

۱۷-۴-۱ پارامترهای گرایش به مرکز

میانگین، میانه و نما پارامترهائی هستند که نشان می‌دهند داده‌های هیدرولوژیکی چگونه و تا چه اندازه حول یک نقطه مرکزی قرار گرفته‌اند. فرض کنید داده‌های سیل یک رودخانه را که در طول دوره آماری اندازه‌گیری و ثبت کرده‌ایم در دسته‌های ۵-۰، ۱۰-۵ و ۱۵-۱۰ و ... متر مکعب در ثانیه قرار داده و تعداد سیل‌هائی را که در هر دسته قرار می‌گیرند شمرده‌ایم. چنانچه هر دسته را با x_i و تعداد سیل موجود در آن دسته را y در نظر گرفته و در یک دستگاه محور مختصات بافتنگار (histogram) آن را رسم کنیم ممکن است شکلی مانند ۱۷-۹ بدست آید که نشان دهنده چگونگی توزیع داده‌هاست. در این شکل یک نقطه مرکزی که بیشترین وقوع در آن اتفاق افتاده است وجود داشته و در ضمن مشخص می‌کند که سایر داده‌ها چگونه از آن دور بوده‌اند.



شکل ۱۷-۹ چگونگی توزیع داده‌های اندازه‌گیری شده از یک پدیده هیدرولوژیکی

الف - میانگین میانگین مقدار متوسطی است که داده‌ها حول آن یکنواخت تقسیم شده‌اند. به مفهوم ریاضی اگر اولین گشتاور را از مبداء بگیریم فاصله میانگین از مبداء به دست می‌آید. یعنی میانگین نقطه‌ای است که اگر گشتاور را حول آن حساب کنیم نقاط منفی ومثبت متعادل شده و مجموع انحرافات از آن صفر می‌شود. در هیدرولوژی میانگین به چند طریق محاسبه می‌شود که عبارتند از:

میانگین ریاضی (حسابی) - میانگین، تخمینی از متغیر است که احتمال وقوع آن درآینده بیشتر از هر مقدار دیگر است. مثلاً اگر میانگین بارندگی سالانه یک ایستگاه بر اساس داده‌های موجود ۲۵۰ میلی متر باشد، رقمی که برای بارندگی سالهای بعد می‌توان پیش بینی کرد و احتمال وقوع آن بیش از هر مقدار دیگر بارندگی‌هاست ، همان ۲۵۰ میلی متر می‌باشد. میانگین ریاضی

(arithmatic mean) را می‌توان از فرمول زیر محاسبه کرد.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (۱-۱۷)$$

در این فرمول $\bar{x}$ میانگین ریاضی، n تعداد داده‌ها، X_i مقدار هر کدام از داده‌ها و i نمایه مربوط به هر کدام از داده‌هاست.

● مثال ۱۷-۱

مقدار بارندگی ماه فروردین طی سالهای مختلف آماری به ترتیب ۱۵، ۳۷، ۲۴، ۲۵، ۲۱، ۲۸ و ۲۵ میلی متر بوده است. میانگین ریاضی (حسابی) بارندگی فروردین ماه چقدر است؟

حل

$$\bar{x} = \frac{15 + 37 + 24 + 25 + 21 + 28 + 25}{7} = 25 \text{ mm}$$

میانگین‌گیری به روش ریاضی دارای محاسن و معایبی است که از آن جمله می‌توان موارد زیر را نام برد.

- محاسبات بسیار ساده است.
- تمام داده‌های آماری در آن مورد استفاده قرار می‌گیرد.
- در هر شرایطی قابل استفاده است. ولی:
- در صورت اشتباه، پیدا کردن نقطهٔ اشتباه دشوار است.
- حتی اگر از یک رقم استفاده نشود اثر زیادی در میانگین دارد.
- وجود ارقام بسیار بزرگ و بسیار کوچک (دو انتها) اثر فراوانی بر نتیجه دارند.

میانگین وزنی - در روش میانگین‌گیری حسابی یا ریاضی به تمامی اعداد یک نوع ارزش داده می‌شود حال آن که در روش وزنی (weighted mean) برای هر عدد وزن و یا ارزش معینی متناسب با اهمیت آن در نظر گرفته می‌شود. اگر هر یک از متغیرها X_1 ، X_2 ،... و X_n و وزنی که به آنها اختصاص داده می‌شود W_1 ، W_2 ،... و W_n باشد میانگین وزنی عبارت است از:

$$\bar{X} = \frac{\Sigma X_i.W_i}{\Sigma W_i} \qquad (۲-۱۷)$$

● مثال ۱۷-۲

تعداد ۷ باران سنج را در یک منطقه نصب کرده‌ایم و هر کدام معرف مساحتی که ذکر شده است می‌باشند. میانگین وزنی بارندگی منطقه را حساب نموده و آن را با میانگین حسابی مقایسه کنید. وزن مربوط به هر یک از ایستگاهها برابر مساحت مربوطه در نظر گرفته شود.

بارانسنج	بارندگی روزانه	مساحت (km^2)
A	35	300
B	30	200
C	30	150
D	25	150
E	30	200
F	35	300
G	25	100

حل

$$\text{میانگین وزنی} = \frac{(35 \times 300) + (30 \times 200) + (30 \times 150) + \dots + (25 \times 100)}{300 + 200 + \dots + 100}$$

$$\text{میانگین وزنی} = 31.25\ \text{mm}$$

اگر به طریق حسابی میانگین را به دست آوریم رقمی معادل ۳۰ میلی متر بدست می‌آید. روش میانگین‌گیری وزنی در تخمین بارندگی یک منطقه به روش تیسن و خطوط همباران کاربرد دارد.

میانگین هندسی در این روش اگر n متغیر داشته باشیم ریشهٔ n ام حاصلضرب n متغیر X_1، X_2، ... و X_n محاسبه می‌شود.

$$\bar{X} = \text{میانگین هندسی} = [(X_1)(X_2)(X_3)\dots(X_n)]^{1/n} \qquad (۳-۱۷)$$

● **مثال ۱۷-۳**

مقدار دبی یک رودخانه را ۵ بار اندازه‌گیری کرده و هر بار اعداد ۱۳۳، ۱۴۱، ۱۲۵، ۱۷۳ و ۱۸۲ متر مکعب در ثانیه به دست آمده است میانگین دبی را به طریق هندسی به دست آورید.

حل

$$\bar{X} = (182 \times 173 \times 125 \times 141 \times 133)^{1/5} = 149\ m^3/sec$$

یک نمونه از روش میانگین‌گیری هندسی (geometric mean) را قبلاً در بخش نفوذ برای پیدا کردن معادلهٔ نفوذ دیدیم. میانگین‌گیری به روش هندسی دارای محاسنی است که از آن جمله می‌توان موارد زیر را نام برد.

- از تمام داده‌ها استفاده می‌شود.
- محاسبات ساده است.
- ارقام اکستریم یا حد انتهایی (اعداد بزرگ و کوچک دو انتها) اثر چندانی در نتیجه ندارند.
- به ارقام بزرگ ارزش زیادی داده نمی‌شود و برعکس به ارقام کوچک بیشتر ارزش داده می‌شود.

میانگین‌گیری به روش هندسی یک ضعف عمده داشته و آن این است که اگر رقمی معادل صفر یا منفی داشته باشیم نمی‌توانیم از این روش استفاده کنیم.

میانگین هارمونیک در این روش، میانگین از فرمول زیر محاسبه می‌شود.

$$\bar{X}(\text{هارمونیک}) = \frac{n}{\frac{1}{X_1} + \frac{1}{X_2} + \ldots + \frac{1}{X_n}} \qquad (۴-۱۷)$$

روش هارمونیک (harmonic mean) را به صورت وزنی نیز می‌توان به کار برد بطوری که اگر X_1، X_2، X_3، ... و X_n مقادیر متغیر و W_1، W_2، ... و W_n وزن اختصاص داده شده به هر کدام باشد میانگین هارمونیک عبارت است از:

$$\bar{X} = \frac{\Sigma W}{\Sigma W/X} \qquad (۵-۱۷)$$

■ **تمرین ۱۷-۱**

در یک حوضه ۹ ایستگاه باران سنجی وجود دارد که منطقه تحت تاثیر هر کدام از آن ها متغیر است. در یک روز بارانی داده‌های اندازه‌گیری شده در هر ایستگاه مطابق جدول زیر بوده است. میانگین ریاضی، هندسی، هارمونیک و میانگین وزنی بارندگی آن روز در حوضه چقدر بوده است.

ایستگاه	A	B	C	D	E	F	G	H	I
بارندگی(mm)	51.8	32	28.7	43.4	38.6	50.5	59.6	31.5	31.7
منطقه تاثیر (Km2)	31.0	58.0	31.5	31.0	86.0	71.0	27.0	43.5	66.0

جواب = ۴۰/۸۷ ، ۳۹/۶ ، ۳۸/۴۲ ، ۳۹/۷۵ متر

برای داده‌های گروه‌بندی شده نیز می‌توان میانگین را محاسبه نمود که نمونه‌ای از این محاسبات در مثال ۱۷-۵ نشان داده شده است.

چنانچه مقدار متوسط هر گروه یا کلاس x_i و فراوانی آن f_i و تعداد کل داده‌ها N باشد میانگین عبارت خواهد بود از:

$$\text{mean} = \frac{\Sigma f_i x_i}{N} \qquad (۶-۱۷)$$

میانگین در پارامترهای برداری برخی پارامترهای هیدرولوژی مانند باد از نوع برداری هستند که هم دارای مقدار و هم دارای جهت می‌باشند. در این پارامترها علاوه بر میانگین ریاضی که از

مقدار عددی آنها حاصل می‌شود جهت غالب متوسط نیز تعیین می‌شود. مثلاً اگر سرعت باد (متر در ثانیه) در جهات مختلف ۸گانه در اختیار باشد، ابتدا مقادیر سرعت باد را در جهات افقی (شرقی) و عمودی (شمالی) تصویرکرده سپس با توجه به ضریب تصویر برای هر یک از جهات مطابق آنچه در فصل ۳گفته شد مقدار میانگین برداری سرعت باد از فرمول زیر محاسبه‌می‌شود.

$$\text{میانگین‌برداری} = [(\text{جمع تصاویرسرعتهاروی‌محورعمودی})^2 + (\text{جمع تصاویرسرعتهاروی‌محورافقی})^2]^{0.5}$$

ب - میانه میانه (median) عبارت است از عددی که در وسط یک سری ارقام قرارگرفته باشد، به شرط آن که اعداد به ترتیب صعودی یا نزولی ردیف شده باشند. مثلاً در مجموعهٔ ۳، ۵، ۶، ۱۰، ۱۲، ۱۴، ۱۷ عدد ۱۰ میانه است. چنانچه یک سری عدد به تعداد n داشته باشیم و n نیز زوج باشد عدد ردیف $\frac{n}{2}$ و $\frac{n}{2}+1$ را متوسط می‌گیریم تا میانه بدست آید. اگر n فرد باشد عدد ردیف $\frac{n+1}{2}$ برابر میانه است. به عبارت دیگر میانه عددی است که ۵۰ درصد داده‌ها بالاتر از آن و ۵۰ درصد داده‌ها پائین‌تر از آن قرار گرفته باشند.

● **مثال ۱۷-۴**

برای اعداد زیر، میانه را بدست آورید

(الف) 12, 11, 7, 5, 8, 14, 9, 13, 16 m3/sec

(ب) 21, 16, 22, 17, 23, 20 m3/sec

حل

پس از ردیف کردن اعداد (n = 9)، میانه عبارت است از:

(الف) 5, 7, 8, 9, 11, 12, 13, 14, 16

$$\text{شماره ردیف میانه} = \frac{9+1}{2} = 5$$

که عدد ردیف ۵ برابر ۱۱ است لذا: $\text{میانه} = 11 \quad m^3/sec$

اعداد سری (ب) عبارتند از: (ب) 16, 17, 20, 21, 22, 23

چون تعداد داده‌ها زوج می‌باشد عدد ردیف $\frac{n}{2}$ برابر ۲۰ و عدد ردیف $\frac{n}{2}+1$ برابر ۲۱ می‌باشد، لذا میانه عبارت است $۲۰/۵ = \frac{۲۰+۲۱}{۲}$.

برای داده‌های گروه‌بندی شده می‌توان میانه را از روی فرمول زیر محاسبه کرد (به مثال ۱۷-۵ توجه شود)

$$\text{Median} = M_d = L_1 + h\,\frac{(N/2) - S_{bm}}{f_{median}} \qquad (۱۷-۷)$$

که در آن:

L_1 = مرز پائین کلاسی که میانه در آن قرار گرفته است.

h = اندازه یا پهنای کلاسی که میانه در آن قرار دارد.

N = تعداد داده‌ها (فراوانی تجمعی)

S_{bm} = جمع تمام فراوانی‌هائی که در زیر کلاس میانه قرار دارند.

f_{median} = فراوانی در کلاسی که میانه در آن قرار دارد.

طرز استفاده از این معادله در مثال ۱۷-۵ نشان داده شده است.

ج - نما نما (mode) متغیری است که دارای بیشترین فراوانی وقوع است. برای متغیرهای پیوسته نما عبارت است از حداکثر مقدار در توزیع فراوانی ولی نما معمولاً برای داده‌های گروهی مورد استفاده قرار می‌گیرد که مقدار آن از فرمول زیر بدست می‌آید:

$$\text{Mode} = L + \left[\frac{(f_1 - f_0)}{(f_1 + f_0) + (f_1 - f_2)}\right] h \qquad (\text{۸-۱۷})$$

در این معادله:

L = حد پائین کلاسی که مد در آن قرار گرفته است.

f_1 = تعداد وقوع در کلاسی که مد قرار دارد.

f_0 = تعداد وقوع در کلاس قبل از مد

f_2 = تعداد وقوع در کلاس بعد از مد

h = فاصله (بازه) کلاس مد

● **مثال ۱۷-۵**

در یک حوضه آبریز ۵۵ روز بارندگی در طول سال وجود داشته که با یک باران سنج اندازه‌گیری شده است. مقادیر داده‌های بارندگی در ۸ دسته مطابق جدول ۱۷-۲ گروه‌بندی شده‌اند. حساب کنید میانگین، میانه و نما (مد) را برای این داده‌ها.

جدول ۱۷-۲

شماره	1	2	3	4	5	6	7	8
کلاس داده ها (mm)	4-50	51-60	61-70	71-80	81-90	91-100	101-110	111-120
تعداد (فراوانی) وقوع	5	8	10	11	12	4	3	2

حل:

برای حل این مساله جدول ۱۷-۳ را تشکیل می‌دهیم:

جدول ۱۷-۳

شماره گروه	کلاس بارندگی	متوسط کلاس (x_i)	فراوانی در کلاس (f_i)	فراوانی تجمعی	$f_i X_i$	ملاحظات
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	41-50	45	5	5	225	*با توجه به تعداد داده‌ها که ۵۵
2	51-60	55	8	13	440	است میانه ۲۷/۵=$\frac{۵۵}{۲}$ می‌باشد که این
3	61-70	65	10	23	650	عدد در گروه یا کلاس چهارم
4	71-80	75	11	34*	825	قرارمی‌گیرد.
5	81-90	85	12**	46	1020	** نما مربوط به کلاس
6	91-100	95	4	50	380	یا گروهی است که بالاترین
7	101-110	105	3	53	315	فراوانی را دارا می‌باشد که
8	111-120	115	2	55	230	در این مثال کلاس پنجم است.
	جمع	=	55		4085	

الف - میانگین برای داده‌های گروهی

$$\text{mean} = \frac{\Sigma f_i f_x}{N} = \frac{4085}{55} = 74.22 \text{ mm}$$

ب - میانه برای داده‌های گروهی

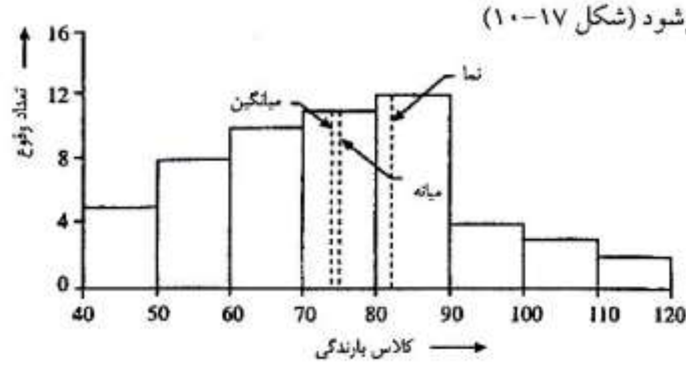
$$\text{Median} = M_d = L_1 + h\,\frac{(N/2) - S_{bm}}{f_{median}} = 71 + (10)\,\frac{(55/2) - 23}{11} = 75.09 \text{ mm}$$

ج - نما برای داده‌های گروهی

$$\text{Mode} = L + \left[\frac{(f_1 - f_0)}{(f_1 - f_0) + (f_1 - f_2)}\right] h$$

$$\text{Mode} = 81 + \left[\frac{(12 - 11)}{\{(12 - 11) + (12 - 41)\}}\right] \times 10 = 82.11 \text{ mm}$$

چنانچه بافتنگار (گراف توزیع) این داده‌ها را رسم کنیم موقعیت میانگین، میانه و نما در آن مشاهده می‌شود (شکل ۱۷-۱۰)



شکل ۱۷-۱۰

به تجربه دریافت شده است که بین میانگین (mean) و میانه (median) و نما (mode) روابط تجربی زیر وجود دارد که می‌توان برای تخمین‌های هیدرولوژیکی از آن استفاده کرد.

$$\text{Mean} - \text{Mode} = 3\,(\text{Mean} - \text{Median}) \qquad (۹-۱۷)$$

$$\text{Mode} = \text{Mean} - 3\,(\text{Mean} - \text{Median}) \qquad (۱۰-۱۷)$$

۱۷-۴-۲ پارامترهای خصوصیات پراکندگی داده‌ها

الف – واریانس

متغیر بودن (variability) داده‌ها از روی پارامتری بنام واریانس (variance) سنجیده می‌شود. به مفهوم ریاضی واریانس، گشتاور دوم داده‌ها حول میانگین می‌باشد که از روی فرمول‌های زیر محاسبه می‌شود.

$$S_n^2 = \frac{1}{n}\Sigma\,(x_i - x_{av})^2 \qquad (\text{۱۱-۱۷ الف})$$

$$S_n^2 = [\,\Sigma\, x_i^2 - \frac{(\Sigma\, x)^2}{n}] \qquad (\text{۱۱-۱۷ ب})$$

در این معادله‌ها n تعداد داده‌ها، x_i مقدار هر یک از داده‌ها x_{av} متوسط داده‌ها و S_n واریانس می‌باشد. چنانچه تعداد داده‌ها از ۳۰ کمتر باشد واریانس محاسبه شده با فرمول‌های بالا نمایه خوبی برای سنجش میزان تغییر داده‌ها نبوده و به اصلاح آماری یک تخمین اریب (biased) از متغیر بودن داده‌ها بدست می‌آید. برای رفع این مشکل فرمول ۱۷-۱۱ را در $\frac{n}{n-1}$ ضرب می‌کنند تا یک تخمین ناآریب از درجه متغیر بودن داده‌ها بدست آید. که آن را با علامت S_{n-1}^2 نشان می‌دهند. لذا خواهیم داشت:

$$S_{n-1}^2 = \frac{1}{n-1}\Sigma\,(x_i - x_{av})^2 \qquad (۱۲-۱۷)$$

برای داده‌های گروه‌بندی شده اگر فراوانی وقوع در هر دسته x_i باشد واریانس عبارت خواهد بود از:

$$S_{n-1}^2 = \frac{1}{n-1}\Sigma\, f_i\,(x_i - x_{av})^2 \qquad (۱۳-۱۷)$$

ب – انحراف از معیار

یکی دیگر از پارامترهائی که متغیر بودن داده‌ها با آن سنجیده می‌شود انحراف از معیار (standard deviation) است که جذر واریانس (با علامت مثبت) می‌باشد و بسته به این که تعداد داده‌ها از ۳۰ بیشتر باشد یا از فرمول:

$$S_n = \sqrt{\frac{1}{n}\,(\Sigma\,(x_i - x_{av})^2} \qquad (۱۴-۱۷)$$

و در مورد تخمین‌های ناآریب از فرمول زیر محاسبه می‌شود:

$$S_{n-1} = \sqrt{\frac{1}{(n-1)} \Sigma(x_i - x_{av})} \qquad (۱۵-۱۷)$$

در بسیاری از محاسبات هیدرولوژیکی معیار ۳۰ - = برای استفاده از فرمول‌های فوق رعایت نمی‌شود و تنها به محاسبه S_n قناعت می‌شود. برخی نیز بجای ۳۰ داده معیار ۵۰ داده را در نظر می‌گیرند.

واریانس و انحراف از معیار معانی مشابه دارند یعنی هر چه مقدار آن‌ها بزرگتر باشد نشان از پراکندگی بیشتر و متغیر بودن زیادتر داده‌ها دارد.

ج - ضریب تغییرات

در پاره‌ای از موارد برای سنجش میزان پراکندگی داده از یک معیار بی‌بعد بنام ضریب تغییرات (coefficient of variation) استفاده می‌شود که عبارت است از نسبت انحراف از معیار به میانگین داده‌ها.

$$C_v = \frac{S}{x_{av}} \qquad (۱۶-۱۷)$$

که در آن C_v ضریب تغییرات و S انحراف از معیار است و بسته به این که تعداد داده‌ها از ۳۰ بیشتر یا کمتر باشد باید بجای S می‌توان در فرمول فوق S_n یا S_{n-1} را قرار داد.

د - خطای معیار

خطای معیار یا برای سنجش درجه اعتبار انحراف از معیار بکار می‌رود (S_e) و یا برای سنجش میانگین (S_{em}). در حالت اول خواهیم داشت:

$$S_e = \frac{S}{\sqrt{2n}} \qquad (۱۷-۱۷)$$

و در حالت دوم:

$$S_{em} = \frac{S}{\sqrt{n}} = \frac{1}{n}\sqrt{\Sigma(x_i - x_{av})^2} \qquad (۱۸-۱۷)$$

در واقع این نمایه انحراف از معیار توزیع آماری پارامترها می‌باشد.

هـ - دامنه

دامنه عبارت اختلاف بین بزرگترین و کوچکترین مقدار داده‌هائی است که به صورت نمونه مورد تجزیه و تحلیل قرار می‌گیرد. مثلاً اگر از حداکثر دبی سیل ۳۰ سال آمار داشته باشیم، اختلاف بین حداکثر و حداقل سیل در این ۳۰ سال دامنه می‌باشد. چنانچه داده‌ها بصورت نرمال بوده و انحراف از معیار آن ها S باشد دامنه را می‌توان از فرمول زیر تخمین زد.

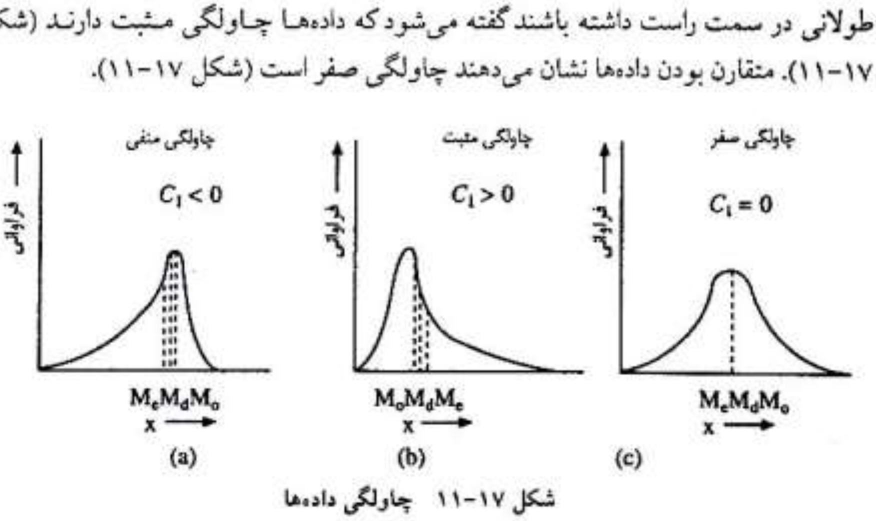
$$R = S\left(\frac{n}{2}\right)^k \qquad (۱۹-۱۷)$$

که n تعداد داده‌ها و k عددی است که بین ۰/۵ تا یک متغیر است ($0.5 < k < 1$). مثلاً اگر ۱۶

سال آمار بارندگی سالانه با انحراف از معیار ۱۰ میلی‌متر داشته باشیم بـه ازاء $k = 0.5$ مـقدار بارندگی ۴۰ و به ازاء $k = 1.0$ مقدار آن ۱۶۰ میلی‌متر است لذا دامنه تغییرات بارندگی در شرایط نرمال باید بین ۴۰ تا ۱۶۰ میلی‌متر باشد.

۱۷-۴-۳ چاولگی

چاولگی یا چولگی (skewness) به مفهوم ریاضی گشتاور سوم داده‌ها حول میانگین است. به زبان ساده چاولگی نشان می‌دهد که آیا داده‌ها حول میانگین متقارن هستند یا خیر. اگر توزیع داده‌ها حول میانگین به نحوی باشد که میانگین به سمت راست تمایل پیدا کرده و یک دنباله طولانی در سمت چپ منحنی توزیع وجود داشته باشد (شکل ۱۷-۱۱) در این صورت گفته می‌شود که چاولگی داده‌ها منفی است و اگر داده‌ها به سمت چپ تمایل داشته و یک دنباله طولانی در سمت راست داشته باشند گفته می‌شود که داده‌هـا چـاولگی مـثبت دارنـد (شکـل ۱۷-۱۱). متقارن بودن داده‌ها نشان می‌دهند چاولگی صفر است (شکل ۱۷-۱۱).



شکل ۱۷-۱۱ چاولگی داده‌ها

چنانچه تعداد داده‌ها از ۳۰ بیشتر باشد چاولگی از فرمول:

$$\alpha\ (\text{یا}\ \mu_3) = \frac{1}{n}\Sigma\,(x_i - x_{av})^3 \qquad (۱۷-۲۰)$$

و اگر تعداد داده‌ها از ۳۰ کمتر باشد سمت راست معادله را در $\frac{n^2}{(n-1)(n-2)}$ ضرب کرده و از معادله زیر استفاده می‌شود.

$$\alpha\ (\text{یا}\ \mu_3) = \frac{n}{(n-1)(n-2)}\Sigma\,(x_i - x_{av})^3 \qquad (۱۷-۲۱)$$

که در آن α یا μ چاولگی، x_i مقدار هر یک از داده‌ها، x_{av} متوسط داده‌ها و n تعداد داده‌ها می‌باشد. با تـوجه بـه مـقدار چـاولگی بـدست آمـده از فـرمول‌های ۱۷-۲۰ و ۱۷-۲۱ در مـحاسبات هیدرولوژیکی از ضریب چاولگی (coefficient of skewness) استفاده می‌شود که عبارت است از:

$$C_s = \frac{\mu_3}{S^3} \qquad (۲۲-۱۷)$$

چنانچه تعداد داده‌ها از ۳۰ تجاوز کند چاولگی را می‌توان از فرمول ساده زیر بدست آورد:

$$S_{ps} = \frac{\text{Mean - Mode}}{\text{انحراف از معیار}} = \frac{3\ (\text{Mean - Median})}{\text{انحراف ازمعیار}} \qquad (۲۳-۱۷)$$

که S_{ps} را چاولگی پیرسون (Pearson's skewness) می‌نامند.

چنانچه چاولگی توزیع داده‌ها به سمت راست باشد $C_s > 0$ و اگر چاولگی توزیع داده‌ها به سمت چپ باشد $C_s < 0$ و درحالتی که توزیع داده‌ها متقارن باشد $C_s = 0$ خواهد بود. در شکل ۱۷-۱۱ موقعیت‌های میانگین (M_e)، میانه (M_d) و نمای (M_o) داده‌ها نیز نشان داده شده است. توجه داشته باشید که اگر برخی داده‌های فرین (اکستریم extreme) در مجموعه داده‌ها وجود داشته باشد گرایش به مرکز پارامترها مانند میانگین بسیار تحت تاثیر قرار خواهد گرفت. تجربه نشان داده است که فرمول ۱۷-۲۳ برای زمانی که تعداد داده‌ها کمتر از ۵۰ باشد زیاد دقیق نخواهد بود و توصیه شده است که در این صورت پس از محاسبه C_s آن را در ضریب $(1 + \frac{8.5}{n})$ ضرب کنیم تا ضریب اصلاح شده چاولگی (C_s) بصورت زیر بدست آید:

$$C'_s = (1 + \frac{8.5}{n})\ C_s \qquad (۲۴-۱۷)$$

۱۷-۴-۴ کشیدگی

برای محاسبه پارامتر کشیدگی (kurtosis) داده‌ها ابتدا گشتاور چهارم (μ_4) داده‌ها حول میانگین محاسبه می‌شود. اگر تعداد داده‌ها از ۳۰ بیشتر باشد فرمول:

$$\mu_4 = (\frac{1}{n})\ (x_i - x_{av})^4 \qquad (۲۵-۱۷)$$

و چنانچه تعداد آن‌ها از ۳۰ کمتر باشد:

$$\mu_4 = [\frac{n^2}{\{(n-1)\ (n-2)\ (n-3)\}}]\ (x_i - x_{av})^4 \qquad (۲۶-۱۷)$$

در این معادله‌ها:

n = تعداد داده‌ها

x_i = مقدار عددی هر یک از داده‌ها

x_{av} = میانگین ریاضی داده‌ها

μ_4 = گشتاور چهارم داده‌ها

پس از محاسبه μ_4 ضریب کشیدگی (kurtosis coefficient) از فرمول زیر بدست می‌آید:

$$\beta_2 = \frac{\mu_4}{S^4} \qquad (۲۷-۱۷)$$

که S انحراف از معیار و β_2 ضریب کشیدگی است. برای سنجش داده‌ها از لحاظ کشیدگی علاوه بر β_2 پارامتری بنام γ_2 نیز محاسبه می‌شود:

$$\gamma_2 = \beta_2 - 3 \qquad (۲۸-۱۷)$$

با توجه به مقادیر β_2 و γ_2 داده‌های هیدرولوژیکی را می‌توان در سه گروه تقسیم‌بندی کرد.

(۱) اگر $\beta_2 > 3$ و $\gamma_2 > 0$ باشد داده‌ها کشیده نام دارند.

(۲) اگر $\beta_2 < 3$ و $\gamma_2 < 0$ باشد داده‌ها پهن نامیده می‌شوند.

(۳) اگر $\beta_2 > 0$ و $\gamma_2 > 0$ باشد داده‌ها نرمال هستند.

● مثال ۱۷-۶

در یک ایستگاه باران سنجی به مدت ۱۰۰ سال داده‌های بارندگی وجود دارد و حداکثر بارش ۲۴ ساعته در طول سال (برحسب میلی‌متر) مطابق جدول زیر می‌باشد. حساب کنید مقادیر میانگین، نمایه، نما، ضریب تغییرات، چاولگی و کشیدگی داده‌ها.

جدول ۱۷-۴ داده‌های حداکثر بارش ۲۴ ساعته در طول سال

77	78	80	93	96	97	98	98	98	97
84	95	83	93	82	91	80	88	84	86
100	100	101	102	106	86	82	87	109	104
75	89	99	96	94	93	92	90	86	78
79	84	83	87	88	89	75	76	76	79
80	81	89	99	104	100	103	104	107	110
110	106	102	107	103	101	101	101	86	94
93	96	97	99	100	102	103	107	107	108
109	94	93	97	98	99	100	97	87	86
94	96	97	98	100	105	106	103	85	84

حل:

با توجه به این که تعداد داده‌ها زیاد می‌باشد و نظر به ارقام و داده‌های جدول ۱۷-۴ که حداقل مقدار اندازه‌گیری شده ۷۵ و حداکثر آن‌ها ۱۱۰ می‌باشد آن‌ها را در ۸ کلاس طبقه‌بندی کرده و جدول ۱۷-۵ را تشکیل داده و محاسبات را به شرح زیر انجام می‌دهیم. کلاس‌های بارندگی عبارتند از: ۷۵ تا ۷۹، ۸۰ تا ۸۴، ۸۵ تا ۸۹، ۹۰ تا ۹۴، ۹۵ تا ۹۹، ۱۰۰ تا ۱۰۴، ۱۰۵ تا ۱۰۹ و ۱۱۰ تا ۱۱۴ میلی متر.

جدول ۱۷-۵ محاسبات آماری داده‌های جدول ۱۷-۴ برای تعیین پارامترهای آماری مورد نیاز مسأله ۱۷-۶

Class	Frequency (f_i)	Relative frequency col. (2)/100	Cumulative relative frequency	Mean class (x_i)	($f_i x_i$) col. (2)× col (5)	x_i^2 col. (5)× col (5)	$f_i(x_i - x_{av})^2$ col. (2)×{(5) −(93.7)}^{2}	$f_i x_i^2$ col. (2)× col (5)^2	$f_i(x_i - x_{av})^3$ col. (2){(5)− −(93.7)}	$f_i(x_i - x_{av})^4$ col. (2){(5)− (93.7)}^{4}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
75–79	09	0.09	0.09	77	693	5929	2510	53361	−41917	7000017
80–84	12	0.12	0.21	82	984	6724	1643	80688	−19219	224866
85–89	15	0.15	0.36	87	1305	7569	673	113535	−4511	30227
90–94	12	0.12	0.48	92	1104	8464	35	101568	−59	100
95–99	19	0.19	0.67	97	1843	9409	207	178771	683	2253
100–104	20	0.20	0.87	102	2040	10404	1378	208080	11436	94917
105–109	11	0.11	0.98	107	1177	11449	1946	125939	25579	344191
110–114	02	0.02	1.00	112	224	12544	670	25088	12257	224303
Sum	100	1.00			9370		9062	887030	−15451	1620874

در جدول ۵-۱۷ ستون اول کلاسهای بارندگی، ستون دوم فراوانی آنها، ستون سوم فراوانی نسبی، ستون چهارم فراوانی نسبی تجمعی و در ستون پنجم متوسط کلاسهای بارندگی ذکر شده است. سایر محاسبات در ستونهای شش الی یازده آورده شده و محاسبات بشرح زیر انجام شده است.

$$\text{میانگین} = \frac{9370}{100}$$

$$\text{میانگین} = 93.70$$

$$\text{میانه} = 95 + 5\,\frac{(100/2) - 48}{19}$$

$$\text{میانه} = 95 + \frac{10}{19}$$

$$\text{میانه} = 95.52$$

$$\text{نما} = 100 + \left\{\frac{5\,(20 - 19)}{(20 - 19) + (20 - 11)}\right\}$$

$$\text{نما} = 100 - \frac{5}{10}$$

$$\text{نما} = 100.50$$

با توجه به این که تعداد داده‌ها ۱۰۰ می‌باشد واریانس عبارت خواهد بود از:

$$\text{واریانس} = \frac{\Sigma f_i x_i^2}{n} - (\frac{\Sigma f_i x_i}{n})^2$$

$$\text{واریانس} = (\frac{887030}{100}) - (\frac{9370}{100})^2$$

$$\text{واریانس} = 8870.30 - (93.7)^2 = 8870.3 - 8779.69$$

$$\text{واریانس} = 90.6$$

به نحو دیگر نیز می‌توان واریانس را محاسبه کرد.

$$\text{واریانس} = \frac{1}{n} \Sigma f_i (x_i - x_{av})^2 = \frac{9062}{100} = 90.6$$

$$\text{انحراف از معیار} = S = (90.6)^{0.5} = 9.52$$

$$\text{ضریب تغییرات} = C_v = \frac{S}{x_{av}} = \frac{9.52}{93.70} = 0.1017 \quad \text{یا}\ 10.17\%$$

$$\text{چاولگی} = \alpha = \frac{1}{n} (x_i - x_{av})^3 = \frac{1}{100} (-15451) = -154.51$$

$$\text{ضریب چاولگی} = \frac{\alpha}{S^3} = \frac{-154.51}{(9.53)^3} = -0.179$$

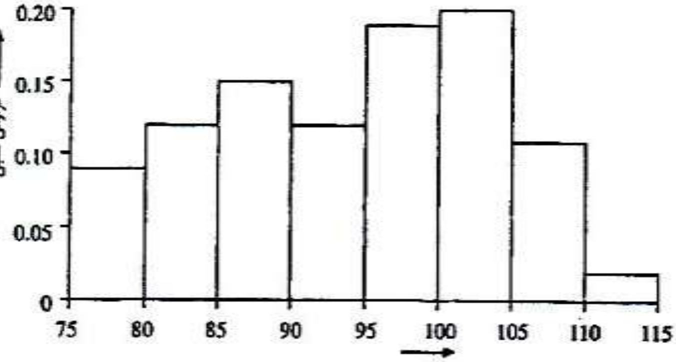
برای محاسبه کشیدگی پارامترهای زیر را محاسبه می‌کنیم.

$$\mu_4 = \frac{1}{n} \Sigma f_i (x_i - x_{av})^4 = 16208.74$$

$$\beta_2 = \frac{\mu_4}{S^4} = \frac{16208.74}{(9.52)^4} = 1.97$$

$$\gamma_2 = \beta_2 - 3 = 1.97 - 3 = -1.03$$

چون $\beta_2 < 3$ و $\gamma_2 < 0$ است لذا داده‌ها به لحاظ کشیدگی پهن و بشقابی شکل می‌باشند. هیستوگرام داده‌ها به صورت زیر می‌باشد.



شکل ۱۷-۱۲

● مثال ۱۷-۷

بارندگی یک ایستگاه در فرودین ماه بمدت ۸ سال پیاپی اندازه‌گیری و به شرح زیر ثبت شده است. مشخصه‌های آماری این داده‌ها را بدست آورید.

$X_5=5$	$X_1=2$ cm
$X_6=3$	$X_2=3$
$X_7=4$	$X_3=4$
$X_8=8$	$X_4=3$

حل

الف - میانگین ریاضی

$$X = (2+3+4+3+5+3+4+8)/8=4.0 \text{ cm}$$

ب - میانگین هندسی

$$G = \sqrt[8]{[(2)(3)(4)(3)(5)(3)(4)(8)]} = 3.69 \text{ cm}$$

ج - میانگین هارمونیک

$$H = 8/(1/2+1/3+1/4+1/3+1/5+1/3+1/4+1/8) = 3.44 \text{ cm}$$

د - میانه

برای بدست آوردن میانه باید داده‌ها را به ترتیب صعودی ردیف کنیم. چون تعداد داده‌ها ۸ می‌باشد پس از ردیف کردن میانه متوسط ارقام ردیف‌های چهارم (۳) و پنجم (۴) یعنی ۳/۵ خواهد بود.

هـ - نما

چون تعداد عدد ۳ در بین داده‌ها بیشتر از بقیه است لذا در این مجموعه نما برابر ۳ می‌باشد.

در این جا مشاهده می‌شود که میانه ۳/۵ و میانگین حسابی ۴ است. بالاتر بودن میانگین حسابی به دلیل این است که داده‌ها دارای چولگی می‌باشند. وجود رقم ۸ در بین داده‌ها باعث شده است که به داده‌ها چولگی مثبت داده شده و میانگین را نسبت به میانه افزایش دهد.

و - دامنه

$$R = 8 - 2 = 6$$

برای بدست آوردن سایر نمایه‌های آماری به شرح جدول زیر عمل می‌شود.

i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
1	$(2-4)=-2$	4	5	$(5-4)=1$	1
2	$(3-4)=-1$	1	6	$(3-4)=-1$	1
3	$(4-4)=0$	0	7	$(4-4)=0$	0
4	$(3-4)=-1$	1	8	$(8-4)=4$	16
					$\Sigma=24$

ز - واریانس (S^2) و انحراف از معیار (S)

$$S^2 = \frac{24}{7} = 3.43 \quad \text{و} \quad S = \sqrt{3.43} = 1.85$$

۱۷-۵ هیستوگرام

در هیدرولوژی برای نشان دادن داده‌ها غالباً از هیستوگرام یا بافتنگار استفاده می‌شود. هیستوگرام یعنی طبقه‌بندی داده‌ها در گروههای مختلف. این که چند گروه تشکیل داده شود و مرز آنها چگونه باشد بسته به نظر هیدرولوژیست بوده و هیچ گونه قانون خاصی در این مورد وجود ندارد. با این وجود دو فرمول پیشنهادی برای تخمین تعداد گروهها در هیستوگرام بکار می‌رود که عبارتند از:

$$m = \sqrt{n} \quad (۲۹-۱۷)$$

$$m = 5 (\log n) \quad (۳۰-۱۷)$$

که n تعداد داده‌ها و m تعداد گروههای هیستوگرام می‌باشد و نتیجه به نزدیک‌ترین عدد صحیح گرد می‌شود. مثلاً اگر تعداد داده‌ها ۲۰۰ باشد مقادیر m از فرمول‌های بالا به ترتیب ۱۴/۱ و ۱۱/۵۱ بدست می‌آید که می‌توانیم یکی از آنها را (۱۴ یا ۱۲) انتخاب کنیم.

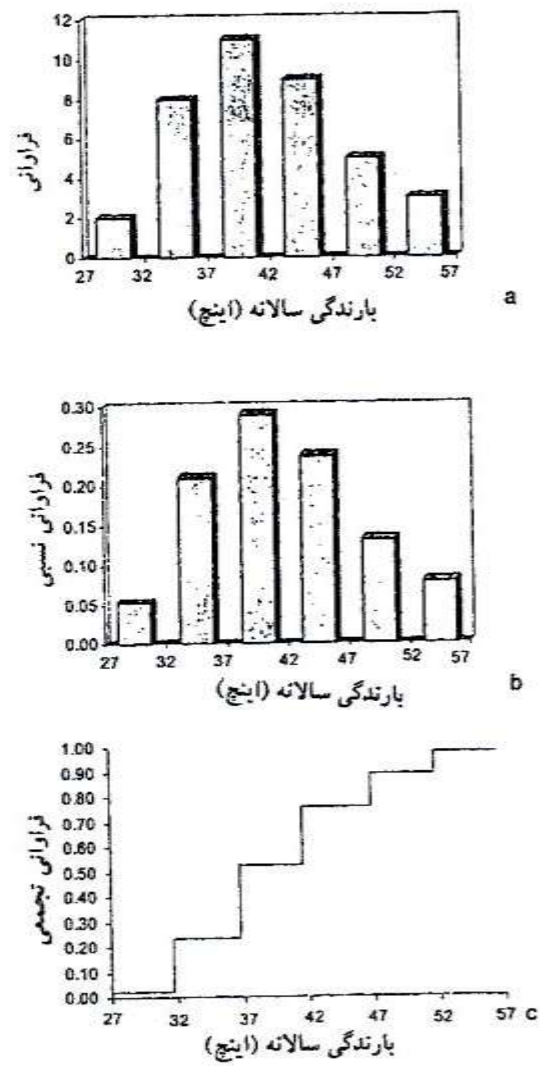
فرض کنید داده‌های جدول ۱۷-۶ مربوط به مقدار بارندگی سالانه (سانتی متر) در یک حوضه آبریز است که هر سال در ایستگاه اندازه‌گیری واقع در حوضه ثبت شده است. وضعیت این داده‌ها را به چند طریق می‌توان بصورت هیستوگرام نشان داد. چون تعداد داده‌ها ۳۸ می‌باشد لذا بر طبق فرمول ۱۷-۲۹ تعداد ۶ گروه برای رسم هیستوگرام توصیه می‌شود البته برای رسم

جدول ۱۷-۶

سال	بارندگی	سال	بارندگی	سال	بارندگی	سال	بارندگی
1950	40.47	1960	41.15	1970	39.14	1980	38.80
	42.06		41.05		47.79		37.38
	45.84		42.63		49.63		40.43
	48.13		34.95		46.06		54.41
	34.04		29.88		37.78		43.66
1955	33.03	1965	29.34	1975	52.13	1985	35.20
	45.90		40.00		33.27		40.42
	32.20		44.82		49.42		33.40
	47.87		35.45		45.95		
	38.37		43.36		52.79		

هیستوگرام مرز یا فاصله گروه‌ها باید طوری انتخاب می‌شوند که هر کدام از اعداد در داخل یکی از گروهها قرار گیرد. برای این منظور ممکن است لازم باشد بر تعداد گروههای هیستوگرام افزوده

یا از آن کسرشود. شکل ۱۷-۱۳ در قسمت a هیستوگرام داده‌ها را بر حسب تعداد مشاهدات نشان می‌دهد. چنانچه تعداد مشاهدات هر گروه را بر تعداد کل مشاهدات (۳۸) تقسیم کنیم می‌توانیم هیستوگرام فراوانی نسبی را مطابق آنچه در قسمت b شکل ۱۷-۱۳ نشان داده شده است بدست آوریم. هیستوگرام فراوانی نسبی می‌تواند تخمین مناسبی از نظر احتمالاتی بدست دهد. مثلاً بر اساس داده‌های مذکور احتمال این که در یک سال بارندگی سالانه در حوضه بین ۳۷ تا ۴۲ سانتی متر باشد ۰/۲۹ یا ۲۹ درصد است. اگر فراوانی‌های نسبی هر گروه را بصورت تجمعی بنویسیم قادر خواهیم بود *فراوانی نسبی تجمعی* (cumulative frequency) را بدست



شکل ۱۷-۱۳ هیستوگرام، فراوانی نسبی و فراوانی تجمعی داده‌های بارندگی

آوریم. با داشتن فراوانی تجمعی می‌توان تخمین زد که احتمال این که مقدار بارندگی در حوضه مساوی یا کمتر از مقدار مشخصی باشد چقدر است. در شکل ۱۷-۱۳ قسمت (c) هیستوگرام فراوانی تجمعی داده‌ها نشان داده شده است. باید توجه داشت که در اینجا به تمام مشاهداتی که در داخل یک گروه قرار می‌گیرند ارزش یکسانی داده شده و برای همه آنها احتمال وقوع یکسان است. ولی روش هائی نیز وجود دارد که هر کدام از داده‌ها به تنهائی مورد بررسی قرار گرفته و احتمال مربوط به آن محاسبه می‌شود. این روش‌ها را ***تحلیل فراوانی‌ها*** (frequeucy analysis) می‌نامند که بعداً در مورد آن بحث خواهد شد.

۱۷-۶ فراوانی وقوع و دوره بازگشت

تجزیه و تحلیل داده‌های هیدرولوژی بدون وارد شدن به محاسبات احتمالاتی امکان‌پذیر نیست. یکی از واژه هایی که در آمار و احتمالات زیاد استعمال می‌شود ***فراوانی وقوع*** یا تعداد دفعاتی است که یک پارامتر مشخص در مدت زمان معین اتفاق می‌افتد.

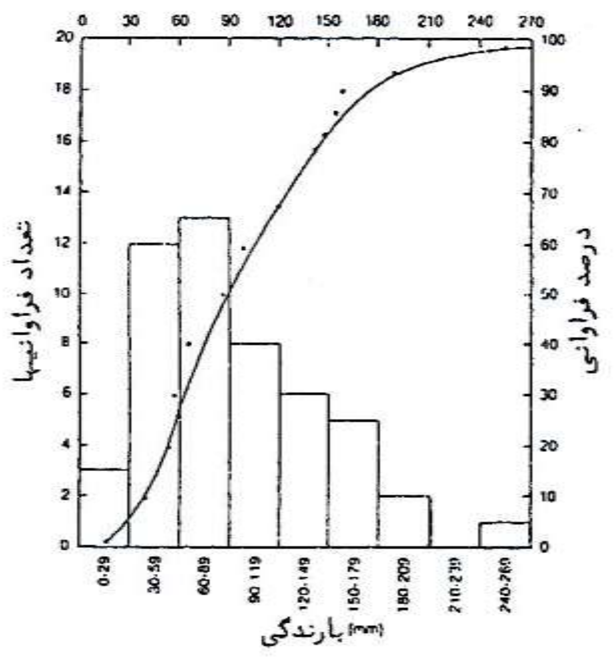
فرض کنید آمار بارندگی فروردین ماه حوضه‌ای را از سال ۱۳۳۷ الی ۱۳۸۶ به مدت ۵۰ سال داشته باشیم و مشاهده شود که متغیرهای این مجموعه آماری بین صفر و ۲۶۹ تغییر می‌کنند. حال اگر رقم صفر تا ۲۶۹ میلی متر را به دسته هایی با فاصله ۲۹ میلی متر تقسیم کنیم تعداد ۹ دسته به دست می‌آید که می‌توانیم جدول فراوانی داده‌ها را تهیه کنیم (جدول ۱۷-۷).
در جدول مذکور در ستون دوم تعداد دفعاتی که بارندگی فروردین ماه در هر دسته قرار می‌گیرد نوشته شده است. مثلاً از ۵۰ سال آماری در ۱۳ مورد بارندگی بین ۶۰ تا ۸۹ میلی متر بوده است و یا در ۲ مورد بین ۱۸۰ تا ۲۰۹ میلی متر بوده و یا در هیچ سالی بارندگی بین ۲۱۰ و ۲۳۹ نبوده است.

اگر فراوانی هر دسته را با دسته‌های ما قبل خود جمع کنیم فراوانی تجمعی حاصل می‌شود، که در ستون سوم جدول مذکور نوشته شده است. ارقام این ستون نشان می‌دهد که مثلاً در ۳۶ مورد از ۵۰ سال آماری بارندگی فروردین ماه از ۱۱۹ میلی متر کمتر بوده است و یا در ۴۹ مورد از ۲۳۹ میلی متر کمتر بارندگی داشته‌ایم. فراوانی وقوع را می‌توان به صورت درصد نیز توصیف کرد که ارقام مربوطه در ستون چهارم جدول نوشته شده‌اند. برطبق ارقام این ستون در ۸۴ درصد موارد بارندگی فروردین ماه حوضه برابر ۱۴۹ میلی متر یا کمتر از آن بوده است.

در هیدرولوژی ساده‌تر خواهد بود که فراوانیهای وقوع یک پارامتر تصادفی را به صورت نمودار نشان دهیم. مثلاً اگر ارقام ستون دوم جدول ۱۷-۷ را به صورت نمودار ستونی رسم کنیم شکلی شکل مشابه ۱۷-۱۴ خواهیم داشت که همان هیستوگرام فراوانی است.

جدول ۱۷-۷ فراوانی وقوع بارندگی فروردین ماه از سال ۱۳۳۷ الی ۱۳۸۶

حدود دسته‌ها	فراوانی (تعداد)	فراوانی تجمعی	فراوانی تجمعی نسبی (٪)
0-29	3	3	6
30-59	12	15	30
60-89	13	28	56
90-119	8	36	72
120-149	5	41	82
150-179	6	47	94
180-209	2	49	98
210-239	0	49	98
240-269	1	50	100



شکل ۱۷-۱۴ هیستوگرام فراوانی وقوع و منحنی احتمال تجمعی

چنانچه این ۵۰ سال آماری را بدون توجه به سال وقوع به ترتیب صعودی ردیف کنیم (جدول ۱۷-۸) احتمال وقوع هر یک از بارندگیها را می‌توانیم از یک فرمول تجربی مانند

$p = \frac{m}{n+1}$ به دست آوریم. در این فرمول P احتمال وقوع، m شمارهٔ ردیف (ستون اول) و n تعداد داده‌هاست که در این مثال ۵۰ می‌باشد. از روی این جدول ملاحظه می‌شود که احتمال این که بارندگی فروردین ماه این حوضه ۱۰۰ میلی متر یا کمتر باشد ۰/۵۸۸ یا ۵۸/۸ درصد است و احتمال این که بارندگی آن از ۱۰۰ میلی متر بیشتر باشد $P = 0.412$ یا ۴۱/۲ درصد خواهد بود و یا شانس این که بارندگی فروردین ماه حوضه بین ۱۰۰ تا ۱۲۵ میلی متر باشد حدود ۰/۱۴ (۱۴ درصد) است (۰/۱۳۷=۰/۵۸۸-۰/۷۲۵).

چنانچه ارقام بارندگی (X) جدول ۱۷-۸ را (ستون‌های دوم و پنجم) نسبت به احتمال وقوع آن (p) که در ستون‌های ۳ و ۶ آمده است، در یک دستگاه محور مختصات رسم کنیم منحنی حاصله را ***توزیع احتمال تجمعی*** (cumulative distribution function, CDF) وقوع بارندگی گویند. مثلاً مطابق این شکل در ۵۰ درصد مشاهدات بارندگی کمتر از ۹۰ میلی متر و در ۳۸ درصد موارد بارندگیها مقدارشان بین ۶۰ تا ۱۲۰ میلی متر است (در شکل ۱۲۰ میلی متر مربوط به احتمال ۰/۶۲ = ۰/۳۸-۱ است).

در مطالب فوق گفته شد که برای محاسبهٔ احتمال داده‌های تجربی از فرمول $p = \frac{m}{n+1}$ استفاده می‌شود. متخصصان هیدرولوژی معتقدند که وقوع پارامترهای آب- هواشناسی مانند بارندگی، سیل، دما و غیره که بصورت تصادفی اتفاق می‌افتند در همه موارد و در همه جا از یک قانون تجربی احتمالاتی خاص مانند $P = \frac{m}{n+1}$ تبعیت نمی‌کنند، بلکه ممکن است این فرمول برای بعضی موارد صادق نباشد. لذاعلاوه بر فرمول فوق فرمول‌های مختلفی دیگری را نیز پیشنهاد نموده‌اند که از جمله آنها می‌توان به معادلات زیر اشاره کرد.

$P = \frac{m}{n}$ (California) فرمول کالیفرنیا (۳۱-۱۷)

$P = \frac{m}{n+1}$ (Weibull) فرمول ویبول (۳۲-۱۷)

$P = \frac{2m-1}{2n}$ (Hazen) فرمول هیزن (۳۳-۱۷)

$P = \frac{m-0.3}{n+0.4}$ (Chegodayev) فرمول چگودیف (۳۴-۱۷)

$P = \frac{3m-1}{3n+1}$ (Tukey) فرمول توکی (۳۵-۱۷)

$P = \frac{m-0.375}{n+0.25}$ (Blom) فرمول بلوم (۳۶-۱۷)

$p = \frac{m-0.44}{n+0.12}$ (Gringorten) فرمول گرینگورتن (۳۷-۱۷)

جدول ۱۷-۸ توزیع تجمعی درصد فراوانی

m	x	$P = \frac{m}{n+1}$	m	x	$P = \frac{m}{n+1}$
1	18	0.019	26	85	0.509
2	23	0.039	27	86	0.529
3	26	0.059	28	89	0.549
4	37	0.078	29	90	0.568
5	41	0.098	30	100	0.588
6	44	0.117	31	105	0.607
7	44	0.137	32	105	0.627
8	48	0.156	33	107	0.647
9	49	0.176	34	108	0.666
10	52	0.196	35	110	0.686
11	54	0.215	36	119	0.705
12	55	0.235	37	125	0.725
13	55	0.254	38	126	0.745
14	57	0.274	39	133	0.764
15	58	0.294	40	144	0.784
16	62	0.313	41	147	0.803
17	62	0.333	42	148	0.823
18	67	0.352	43	153	0.843
19	67	0.372	44	154	0.862
20	68	0.392	45	157	0.882
21	71	0.411	46	158	0.901
22	72	0.431	47	171	0.921
23	78	0.450	48	189	0.941
24	78	0.470	49	206	0.960
25	79	0.490	50	247	0.980

در هیدرولوژی غالباً بجای کلمه احتمال از واژه دوره ***بازگشت*** (return period) نیز استفاده می‌شود. دورهٔ بازگشت عکس احتمال است و آن تعداد سالهایی است که بطور متوسط بین وقوع دو حادثه مشابه وجود دارد. اگر دورهٔ بازگشت T (سال) و احتمال وقوع P باشد خواهیم داشت.

$$T = \frac{1}{P} \qquad (۳۸-۱۷)$$

مثلاً احتمال این که بارندگی فروردین ماه حوضهٔ در مثال قبل معادل ۵۵ میلی متر یاکمتر باشد ۲۵ درصد است (۰/۲۵) لذا دورهٔ برگشت آن $\frac{۱}{۰/۲۵}$ یا ۴ سال است یعنی چنین حالتی بطور متوسط هر ۴ سال یک بار اتفاق می‌افتد. با این تعریف؛

- احتمال ۱۰ درصد یعنی دورهٔ برگشت ۱۰ سال (هر ۱۰ سال یکبار)
- احتمال ۰/۵ درصد یعنی دورهٔ برگشت ۲۰۰ سال (هر ۲۰۰ سال یکبار)
- احتمال ۰/۹۸ یا ۹۸ درصد یعنی دورهٔ برگشت ۱/۰۲ سال (در هر ۱۰۲ سال ۱۰۰ بار)
- احتمال ۰/۹۹ یا ۹۹ درصد یعنی دورهٔ برگشت ۱/۰۱ سال (در هر ۱۰۱ سال ۱۰۰ بار)

لذاباران ۱۰۰ ساله بارانی است که احتمال وقوع باران مشابه با آن یا بیشتر از آن یک درصد باشد.

۱۷-۷ توابع توزیع احتمال

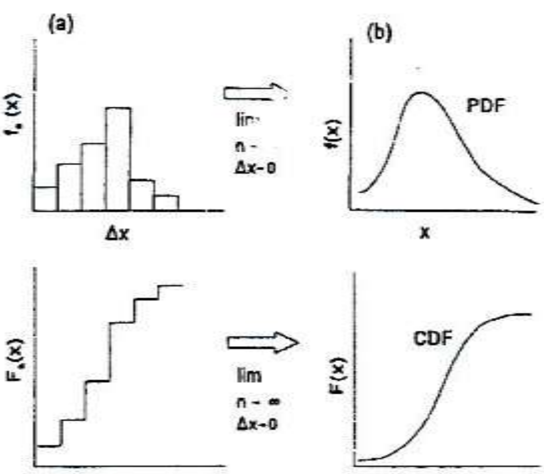
اگر چنانچه تعداد داده‌های هیدرولوژی n بوده و آن‌ها را جهت رسم هیستوگرام دسته بندی کرده باشیم و فرض کنیم که در هر دسته تعداد n_i داده قرار گیرد لذا فروانی نسبی در هر دسته n_i/n خواهد بود. بنابراین ***تابع فراوانی نسبی*** (relative frequency function) عبارت خواهد بود از:

$$f(x_i)=\frac{n_i}{n} \qquad (۱۷-۳۹)$$

از جمع کردن مقادیر فراوانی‌های نسبی با همدیگر ***تابع فراوانی تجمعی*** بدست می‌آید (cumulative frequency funtion) که با علامت $F(x_i)$ نشان داده می‌شود.

$$F(x_i) = \Sigma f(x_i) \qquad (۱۷-۴۰)$$

در شکل ۱۷-۱۵ توابع فراوانی نسبی و فراوانی تجمعی بصورت هیستوگرام در سمت چپ رسم شده‌اند این هیستوگرام‌ها بر اساس داده‌های گسسته می‌باشند.توابع فراوانی نسبی و فراوانی تجمعی که در بالا گفته شد برای تعداد محدودی از داده‌های نمونه‌گیری شده می‌باشد که اگر بخواهیم آن را برای کل جمعیت آماری تعمیم دهیم در این صورت باید داده‌های گسسته را به داده‌های پیوسته تبدیل کرد تا منحنی‌های پیوسته‌ای مشابه آنچه در شکل ۱۷-۱۵ نشان داده شده است بدست آید. لذا بجای ***تابع فراوانی نسبی*** از probability density function, PDF یا ***تابع چگالی احتمال*** وبجای تابع ***فراوانی‌تجمعی*** ازنام cumulative distribution function, CDF یا تابع چگالی تجمعی استفاده می‌شود که به ترتیب با علائم PDF و CDF نشان داده می‌شوند.



شکل ۱۷-۱۵ تبدیل توابع نمونه‌های گسسته به توابع پیوسته

در هیدرولوژی سعی می‌شود برای داده‌ها توابع احتمالاتی مناسبی پیدا و رسم شود تا از روی آن‌ها بتوان مقدار متغیر مورد نظر را به ازاء احتمالات مختلف محاسبه کرد.

۱۷-۸ توزیع‌های احتمالاتی تئوری

پارامترهای آماری که در بخش‌های قبل تشریح شد مشخص کننده توزیع آماری نمونه‌هائی است که بعنوان داده‌های هیدرولوژیکی در اختیار می‌باشد. حال یک نفر هیدرولوژیست باید بتواند از این پارامترها استفاده کرده و با انتخاب یک توزیع احتمالاتی وقایع هیدرولوژیکی را برای آینده پیش‌بینی نماید. این کار به ما کمک خواهد کرد که بتوانیم بگوئیم در طول عمر مفید یک پروژه مثلاً یک سیل با مقدار مشخص چند بار و با چه احتمالی بوقوع خواهد پیوست. با برازش دادن داده‌های هیدرولوژیکی به یک توزیع احتمالاتی می‌توان احتمال وقوع یک پدیده تصادفی را مشخص کرد.

در هیدرولوژی تعداد زیادی توزیع‌های فراوانی وجود دارند که به صورت‌های گوناگون با برازش داده‌ها بر آن ها مورد استفاده قرار می‌گیرند. برخی از این توزیع‌ها معمولاً به همان صورت تئوری استفاده می‌شوند که در زیر به تشریح برخی از آن‌ها که کاربرد زیادتری دارند می‌پردازیم:

۱۷-۸-۱ توزیع‌های گسسته

(۱) توزیع احتمالاتی دو جمله‌ای

توزیع دوجمله‌ای (binomial) فقط برای تجزیه و تحلیل داده‌های گسسته مورداستفاده قرار می‌گیرد مانند پیش‌بینی تعداد روزهای بارانی، تعداد روزهای یخبندان و یا تعیین احتمال وقوع یا عدم وقوع یک حادثه. در این توزیع یا یک واقعه اتفاق می‌افتد که احتمال آن P است و یا این که رخ نخواهد داد که احتمال آن را با q نشان می‌دهیم و مسلم است که $q = 1 - p$ خواهد بود، زیرا بین این دو حالت دیگری وجود ندارد. اگر تعداد داده‌های موجود n باشد شانس این که واقعه‌ای با مشخصه x رخ دهد، بر اساس توزیع دو جمله‌ای به لحاظ تئوری عبارت خواهد بود از:

$$P(x) = \left\{\frac{n!}{[x!\ (n - x)!]}\right\} P^{x}\, q^{\,n-x} \qquad (۱۷-۴۱)$$

توصیف اجزاء به این فرمول در دو مثال زیر روشن شده است.

● **مثال ۱۷-۸**

بر اساس داده‌های موجود از یک محل احتمال این که در روزهای ماه خرداد بارندگی وجود داشته باشد ۲۰ درصد است. قرار است در این محل یک کار ساختمانی انجام شود که لازم است در طول ۵ روز هیچ گونه بارانی وجود نداشته باشد. احتمال این که در طی ۵ روز پیاپی در ماه خرداد باران وجود نداشته باشد چقدر است؟ احتمال این که در طی ۵ روز فقط یک روز بارندگی وجود داشته باشد چقدر خواهد بود؟

حل:

الف – برای آن که در طی ۵ روز هیچ روز بارانی وجود نداشته باشد

$n = 5$

$p = 0.2$

$q = 1 - P = 1 - 0.2 = 0.8$

$x = 0$

$P\ (x = 0) = ?$

$P_{(x=0)} = \{\frac{n!}{[x!\ (n-x)!]}\}\ P^x\ q^{n-x}$

$P_{(x=0)} = \{\frac{5!}{[0!\ (5 - 0)!]}\}\ (0.2)^0\ (0.8)^{5-0}$

$P_{(x=0)} = 0.328$

بنابراین با ۳۲/۸ درصد در ماه خرداد در ۵ روز پیاپی هیچ گونه بارندگی روزانه وجود نخواهد داشت.

ب – برای آن که در طول ۵ روز فقط یک روزبارندگی داشته باشیم

$n = 5$

$P = 0.2$

$q = 0.8$

$x = 1$

$P_{(x=1)} = ?$

$P_{(x=1)} = \{\frac{5!}{[1!\ (5-1)!]}\}\ (0.2)^1\ (0.8)^{5-1}$

$P_{(x=1)} = 0.41$

لذا احتمال این که در طی ۵ روز یک روز آن بارانی باشد ۴۱ درصد است.

● مثال ۱۷-۹

در طی احداث یک سد که ۱۰ سال به طول خواهد انجامید لازم است که یک فراز بند (coffer dam) برای مهار و انحراف سیل‌ها احداث شود. فراز بند طوری طراحی شده است که سیل‌های با دوره برگشت ۵ ساله را مهار نماید. با چه احتمالی در طول این ۱۰ سال چنین سیل‌هائی اتفاق نخواهد افتاد؟ با چه احتمالی در طول ۱۰ سال کارهای ساختمانی ۲ بار چنین سیل‌هائی اتفاق خواهد افتاد؟

حل:

$$P = \frac{1}{5} = 0.2$$
$$q = 1 - P = 1 - 0.2 = 0.8$$
$$n = 10$$
$$x = 0$$
$$P_{(x=1)} = \{\frac{n!}{[x!\ (n-x)!]}\}\ P^x\ q^{n-x}$$
$$P_{(x=1)} = \{\frac{10!}{[0!\ (10-0)!]}\}\ (0.2)^0\ (0.8)^{10-0} = 0.107$$

لذا احتمال این که در طول ۱۰ سال سیل اتفاق نیتفد ۱۰/۷ درصد است و برای آن که ۲ بار اتفاق افتد:

$$P = 0.2$$
$$q = 0.8$$
$$n = 10$$
$$x = 2$$
$$P_{(x=1)} = \{\frac{10!}{[2!\ (10-2)!]}\}\ (0.2)^2\ (0.8)^{10-2} = 0.302$$

با احتمال ۳۰/۲ درصد در طول ۱۰ ساله کارهای ساختمانی سیل دوبار اتفاق خواهد افتاد.

(۲) توزیع احتمالاتی پواسون

یکی دیگر از توابع توزیع داده‌های گسسته توزیع پواسون (Poisson) است که می‌توان آن را حالت خاصّی از توزیع دو جمله به حساب آورد. توزیع پواسون زمانی در هیدرولوژی بکار می‌رود که:

الف - دوره موردنظر (n) بزرگ باشد.

ب - احتمال وقوع پدیده موردنظر (P) کوچک باشد.

ج - حاصلضرب P و n که با λ نشان داده می‌شود ($\lambda = np$) کوچک باشد.

در این صورت به لحاظ تئوری احتمال وقوع یک حادثه با مشخصه x برابر است با:

$$P\ (x) = \frac{\lambda^x\ e^{-\lambda}}{x!} = \frac{(P.n)^x\ (e)^{-P.n}}{x!} \qquad (۴۲-۱۷)$$

توزیع پواسون در هیدرولوژی معمولاً در مطالعات خشکسالی و سیل که حدود ۱۰۰ سال آمار وجود داشته باشد مورد استفاده قرار می‌گیرد.

● مثال ۱۷-۱۰

در یک منطقه آمار ۸۰ سال بارندگی وجود دارد. بررسی این داده‌ها نشان می‌دهد که احتمال این که مقدار حداکثر بارش ۲۴ ساعته از ۳۰ میلی‌متر تجاوز کند ۰/۰۲ است (۲ درصد). حساب کنید در طی ۱۰ سال آینده احتمال این که ۳ بارش ۲۴ ساعته وجود داشته باشد که مقدار هر یک از آن‌ها از ۳۰ میلی‌متر تجاوز کند چقدر است؟

حل:

$$P = 0.02$$

$$n = 10$$

$$x = 3$$

$$\lambda = P.n = (0.02)\,(10) = 0.2$$

$$P_{(x=3)} = \frac{\lambda^{x}\,.\,e^{-\lambda}}{x!} = \frac{(0.2)^{3}\,(e)^{-0.2}}{3!} = 0.0011$$

بنابراین فقط ۰/۱۱ درصد احتمال خواهد داشت که در طول ۱۰ سال آینده ۳ بارش ۲۴ ساعته بالاتر از ۳۰ میلی‌متر داشته باشیم.

۱۷-۸-۲ توزیع‌های پیوسته

در هیدرولوژی بسیاری از وقایع بخشی از یک فرایند پیوسته می‌باشند. مثلاً سیل را نمی‌توان مثل یک روز بارانی به شمار آورد که بتوان گفت باران هست یا نیست، بلکه سیل دارای مقدار است که می‌تواند از ∞- تا ∞+ عدد داشته باشد. مثلاً در شکل ۱۷-۱۴ اگر بجای ۵۰ عدد هزارها عدد می‌داشتیم و فاصله دسته‌ها را بسیار کوچک می‌گرفتیم هیستوگرام به یک منحنی پیوسته‌ای زنگوله‌ای شکل تبدیل می‌شد. لذا برای این وقایع باید از توزیع‌های پیوسته استفاده نمود. توزیع‌های پیوسته که در هیدرولوژی زیاد مورد استفاده قرار می‌گیرند عبارتند از:

- توزیع احتمالی نرمال (normal)
- توزیع احتمالاتی لوگ – نرمال (log - normal)
- توزیع احتمالاتی گاما (Gamma)
- توزیع احتمالاتی پیرسون تیپ ۳ (Pearson type - III)
- توزیع احتمالاتی لوگ – پیرسون تیپ ۳ (log - Pearson III)
- توزیع احتمالاتی گامبل موسوم به مقادیر اکستریم تیپ ۱ (Gumble)
- توزیع احتمالاتی گامبل موسوم به مقادیر اکستریم تیپ ۳

در این‌جا به‌شرح برخی ازآن‌ها می‌پردازیم وسپس کاربردآن‌ها را مورد بحث قرار خواهیم داد.

(۱) توزیع احتمالاتی نرمال

اگر چنانچه یک متغیر هیدرولوژی x مانند سیل یا بارندگی را اندازه‌گیری کرده و میانگین داده‌ها μ و انحراف از معیار آن‌ها σ باشد تابع چگالی احتمال (PDF) متغیر x که از $-\infty$ تا $+\infty$ تغییر می‌کند زمانی نرمال خواهد بود که به لحاظ تئوری از معادله زیر تبعیت کند.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2\right] \qquad (۴۳-۱۷)$$

چنانچه بجای $\frac{x-\mu}{\sigma}$ مقدار t را که متغیر تصادفی نرمال استاندارد (standard normal variate) نام دارد قرار دهیم معادله فوق بصورت زیر خواهد بود.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) \qquad (۴۴-۱۷)$$

که اگر آن را بین دو حد $-\infty$ تا $+\infty$ انتگرال‌گیری کنیم توزیع نرمال استاندارد (standard normal distribution) بدست می‌آید. از آن جائی که متغیر تصادفی از $-\infty$ تا $+\infty$ تغییر می‌کند، سطح زیر منحنی توزیع نرمال بین دو حد مشخص برابر احتمال وقوع آن حادثه خواهد بود. مثلاً با این توزیع می‌توانیم مشخص کنیم که احتمال آن که سیل بین ۱۰۰ تا ۱۱۰ متر مکعب در ثانیه باشد چقدر است و یا بگوئیم احتمال این که سیل از ۱۱۰ متر مکعب در ثانیه بیشتر باشد چقدر خواهد بود، ولی هیچ وقت نمی‌توانیم بگوئیم احتمال این که سیلی با مقدار دقیقاً ۱۱۰ متر مکعب در ثانیه رخ دهد چقدر است. زیرا سیل یک واقعه پیوسته است و ما از یک توزیع احتمالاتی پیوسته استفاده کرده‌ایم.

همان طور که گفته شد متغیر تصادفی نرمال استاندارد (t) عبارت است از:

$$t = \frac{x-\mu}{\sigma} \qquad (۴۵-۱۷)$$

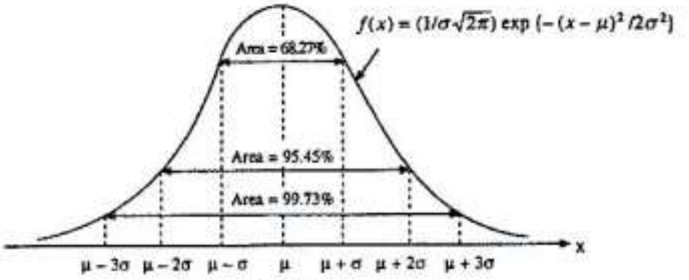
$$x = \mu + \sigma t \qquad (۴۶-۱۷)$$

معادله فوق برای داده‌های از $-\infty$ تا $+\infty$ صادق است حال آن که ما در هیدرولوژی فقط تعداد معدودی داده داریم که میانگین آن‌ها x می‌باشد، لذا معادله بالا بصورت زیر تعدیل می‌شود.

$$x = \bar{x} + \sigma t \qquad (۴۷-۱۷)$$

که به جای σ نیز همان انحراف از معیار نمونه‌ها (S) قرار داده می‌شود. در شکل ۱۷-۱۶ منحنی توزیع نرمال رسم شده و سطح زیر منحنی به ازاء مقادیر σ برابر ±1 ، ±2 و ±3 نشان داده شده است. این بدان معنی است که اگر داده‌ها توزیع نرمال داشته باشند با احتمال ۹۵/۴۵ درصد باید بین دو حد $\bar{x} - 2\sigma$ و $\bar{x} + 2\sigma$ قرار گیرند. برای مقادیر مختلف t سطح زیر منحنی نرمال در جدول ۱۷-۹ داده شده است. طرز استفاده از این جدول به این صورت است که به ازاء مقدار t سطح زیر منحنی که مربوط به یک طرف خط تقارن منحنی است بدست آمده و سپس آن را ضرب در ۲ می‌کنیم تا احتمال محاسبه شود. مثلاً اگر $t = 1.23$ باشد به ازاء $t = 1.23$ سطح زیر منحنی در یک

طرف خط تقارن ۰/۳۹۰۷ و سطح کل ۲ برابر آن یعنی ۰/۷۸۱۴ یا ۷۸/۱۴ درصد می‌باشد.



شکل ۱۷-۱۶ سطح زیر منحنی توزیع احتمال نرمال

منحنی توزیع نرمال برای دوره‌های بالاتر از یک سال مورد استفاده قرار می‌گیرد. یعنی بارش سالانه یا سیل سالانه را می‌توان باتوزیع نرمال برازش داد ولی نباید سعی کرد که داده‌های روزانه، هفته‌ای و ماهانه را با این توزیع برازش داد. طرز برازش داده‌ها را با این توزیع بعداً بحث خواهیم کرد.

جدول ۱۷-۹ سطح زیرمنحنی استاندارد توزیع نرمال برای مقادیر مختلف $[t = (x-\mu)/\sigma]$

t	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0159	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0735
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1271	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1884	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2464	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4657	0.4370	0.4382	0.4394	0.4406	0.4418	0.4430	0.4441
1.6	0.4452	0.4463	0.4474	0.4485	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4762	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4835	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4865	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4990	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4980	0.4980	0.4981
2.9	0.4981	0.4982	0.4983	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4986	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4980	0.4990

معادله تئوری توزیع نرمال مربوط به چگالی احتمال (PDF) می‌باشد و اگر بخواهیم توزیع چگالی تجمعی (CDF) را بدست آوریم باید از آن انتگرال گرفته شود که کاری بسیار مشکل است. برای رفع این مشکل و به نحوی که بتوانیم به ازاء یک احتمال مشخص (P) مقدار متغیر x را بدست آوریم روش‌ها و جداول مختلف توسط متخصصان ارائه شده است تا حل آن را ساده نماید. از جمله این روش‌ها جدول موسوم به تبدیل Z (Z - transform) برای حل معادله توزیع نرمال است. نمایه Z از فرمول زیر بدست می‌آید:

$$Z = \frac{(x_i - \bar{x})}{S} \qquad (۴۸-۱۷)$$

که در آن:

$\bar{x}$= میانگین داده‌ها

S = انحراف از معیار داده‌ها

x_i = مقدار متغیر موردنظر

مقادیر احتمال تجمعی به ازاء مقادیر مختلف Z در جدول ۱۷-۱۰ آورده شده و نحوه استفاده از آن نیز در مثال ۱۷-۱۰ شرح داده شده است.

جدول ۱۷-۱۰ احتمال تجمعی در توزیع نرمال استاندارد

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.60	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.70	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.80	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.90	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.00	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.10	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.20	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.60	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.00	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.20	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.30	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.40	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.50	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.60	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.70	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.80	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.90	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.00	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.10	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.20	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.30	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.40	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

● مثال ۱۷-۱۱

میانگین و انحراف از معیار داده‌های بارندگی سالانه یک حوضه به ترتیب ۴۱/۲۸ و ۶/۴۹ سانتی متر است اگر فرض کنیم توزیع تئوری نرمال با داده‌های مشاهده شده منطبق باشد با استفاده از تابع تئوری نرمال حساب کنید:

الف - با چه احتمالی مقدار بارندگی سالانه از ۴۵ سانتی متر کمتر خواهد بود.

ب - با چه احتمالی مقدار بارندگی سالانه بین ۴۳ و ۴۵ سانتی متر خواهد بود.

ج - با چه احتمالی بارندگی سالانه از ۳۵ سانتی متر کمترخواهد بود.

د - با چه احتمالی دقیقاً در هر سال ۳۵ سانتی متر بارندگی خواهیم داشت.

حل

الف -

$\overline{x} = 41.28$

$S = 6.49$

$x_i = 45$

$$Z = \frac{45 - 41.28}{6.49}$$

$Z = 0.57$

از جدول ۱۷-۱۰ به ازاء $Z = 0.57$ مقدار احتمال تجمعی ۰/۷۱۵۷ بدست می‌آید. لذا با احتمال ۷۱/۵۷ درصد مقدار بارندگی در هر سال کمتر از ۴۵ سانتی متر خواهد بود.

ب - مقادیر Z به ازاء ۴۵ و ۴۳ سانتی متر بارندگی ۰/۵۷ و ۰/۲۷ خواهد بود که مقادیر احتمال متناظر آن‌ها از جدول ۱۷-۱۰ به ترتیب ۰/۷۱۵۷ و ۰/۶۰۶۴ بدست می‌آید لذا احتمال این که بارندگی سالانه بین ۴۵ و ۴۳ سانتی متر باشد تفاصل دو رقم فوق یعنی ۱۰/۹ درصد است. زیرا:

$p(43<x<45) = 71.57 - 60.64$

$p(43<x<45) = 10.9$

ج - به ازاء ۳۵ سانتی متر بارندگی مقدار Z عدد منفی خواهد بود زیرا:

$$Z = \frac{35 - 41.28}{6.49}$$

$Z = -0.97$

چون Z منفی است ابتدا به ازاء ۰/۹۷ مقدار احتمال را بدست آورید (که برابر ۰/۸۳۴۰ خواهد بود). سپس آن را از یک کسر کنید تا احتمال مربوطه محاسبه شود. در این صورت احتمال این که بارندگی سالانه از ۳۵ سانتی متر کمتر باشد ۰/۱۶۶ یا ۱۶/۶ درصد است. زیرا:

$p(x<35) = 1 - 0.834 = 0.166$

د - برای بدست آوردن این که با چه احتمالی بارندگی دقیقاً ۳۵ سانتی متر باشد بهمان

روش بند ب عمل می‌شود که عبارت خواهد بود از:

$p(35<x<35) = 0.166 - 0.166$

$p(35<x<35) = 0$

بنابراین توزیع احتمال نرمال مانند دیگر توزیع‌های پیوسته احتمال این که دقیقاً یک مقدار مشخص اتفاق افتد صفر است زیرا روی منحنی توزیع احتمال بی نهایت نقطه وجود دارد و اگر بخواهیم احتمال وقوع آن را محاسبه کنیم مقدار آن $\frac{1}{\infty}$ خواهد شد که مساوی صفر است ولی در توزیع‌های گسسته امکان این که برای هر متغیر احتمال وقوع آن را حساب کنیم وجود دارد.

(۲) توزیع احتمالاتی لوگ - نرمال

توزیع لوگ - نرمال حالت خاصی از توزیع نرمال است. به این صورت که به جای استفاده از متغیر تصادفی از لگاریتم آن در مبنای e استفاده شده است.در این توزیع از هر یک از متغیرهای x_i لگاریتم گرفته و آن را Z_i می‌نامیم [$Z_i = \ln(x_i)$] که x_i همان متغیر مشاهده شده می‌باشد. سپس μ_z که میانگین Z_iها می‌باشد و σ_z که انحراف از معیارها Z_i هاست محاسبه می‌گردد. چنانچه یک سری از داده‌های x داشته باشیم میانگین (μ_z) و انحراف از معیار (σ_z) برای استفاده در توزیع لوگ - نرمال عبارت خواهد بود از:

$$\mu_x = e^{(\mu_z + \frac{\sigma_z^2}{2})} \qquad (۴۹-۱۷)$$

$$\sigma_x^2 = \mu_x^2 (e^{\sigma_z^2} - 1) \qquad (۵۰-۱۷)$$

و تابع توزیع چگالی احتمال f(x) از معادله زیر تبعیت می‌کند.

$$f(x) = [\frac{1}{x\sigma\sqrt{2\pi}}]\, e^{[-\frac{1}{2}\{\frac{(\ln x - \mu_z)}{\sigma_z}\}^2]} \qquad (۵۱-۱۷)$$

دراین معادله مقادیر $x>0$ بوده و μ_z برابر میانگین zها (Z_{av}) می‌باشد. با داشتن یک سری داده‌های هیدرولوژیکی و محاسبه میانگین و انحراف از معیار آن‌ها بر اساس فرمول‌های ۱۷-۴۹ و ۱۷-۵۰ و با استفاده از جدول ۱۷-۱۰ در واقع داده‌ها با توزیع لوگ - نرمال برازش داده شده‌اند.

(۳) توزیع تئوری گاما

داده‌هائی دارای توزیع گاما هستند که با توجه به مقادیر میانگین (X_{av})، انحراف از معیار (S^2) و ضریب تغییرات (C_v) معادله تابع PDF آن‌ها بصورت زیر باشد:

$$f(x) = \{\frac{\lambda^{\gamma}\ x^{\gamma-1}}{\Gamma(\gamma)!}\}e^{-\lambda x} \qquad (۵۲-۱۷)$$

که در آن $\frac{X_{av}}{\sigma^2}=\lambda$ و $\frac{1}{C_{v^2}}=\gamma$ بوده و Γ علامت تابع گاما (gamma function) می‌باشد. در ریاضیات حل معادلات دیفرانسیلی گاهی اوقات با حاصلضرب‌هائی به شکل $x(x+1)(x+2)\ ...\ (x+n)$ مواجه می‌شویم که این حاصلضرب را می‌توان برحسب تابع گاما $\Gamma(x)$ بیان نمود. تابع گاما برای اعداد مثبت $(x>0)$ به صورت زیر تعریف می‌شود.

$$F(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt \qquad (۵۳-۱۷)$$

که بحث پیرامون آن خارج از موضوع این کتاب می‌باشد. بر اساس قوانین ریاضی در توابع گاما می‌توان نوشت:

$$\Gamma(\gamma+1) = \gamma\Gamma\gamma \qquad (۵۴-۱۷)$$

بنابراین چنانچه در یک مجموعه داده‌های هیدرولوژیکی ضریب تغییرات مثلاً 0.42 باشد خواهیم داشت:

$$C_v = 0.42$$
$$C_v^2 = 0.1764$$
$$1/C_v^2 = 5.67$$
$$\gamma = 5.67 = 4.67 + 1$$

لذا می‌توان نوشت:

$$\begin{aligned}\Gamma 5.67 = \Gamma(4.67+1) &= 4.67\ \Gamma\ 4.2\\ &= 4.67\ (3.67\ \Gamma\ 3.67)\\ &= 4.67 \times 3.67\ (2.67\Gamma\ 2.67)\\ &= 4.67 \times 3.67 \times 2.67\ (1.67\ \Gamma\ 1.67)\end{aligned}$$

بنابراین کافی است تابع گامای ۱/۶۷ را داشته باشیم تا در نهایت تابع گامای γ بدست آید. برای اعداد $n = 1$ تا $n = 2$ مقدار تابع گاما (Γn) از جدول ۱۷-۱۱ قابل استخراج است. در مثال فوق تابع گامای ۱/۶۷ از جدول مذکور برابر ۰/۹۰۳۳۳ بوده لذا:

$$\Gamma\ 5.67 = (4.67)\ (3.67)\ (2.67)\ (1.67)\ (0.90333) = 69.03$$

که می‌توان مقدار فوق را در معادله توزیع گاما قرار داد.

جدول ۱۷-۱۱ تابع گاما برای n (Γn)

n	$\Gamma(n)$	n	$\Gamma(n)$	n	$\Gamma(n)$
1.00	1.000000	1.34	0.892216	1.68	0.905001
1.02	0.988844	1.36	0.890185	1.70	0.908639
1.04	0.978438	1.38	0.888537	1.72	0.912581
1.06	0.968744	1.40	0.887264	1.74	0.916826
1.08	0.959725	1.42	0.886356	1.76	0.921375
1.10	0.951351	1.44	0.885805	1.78	0.926227
1.12	0.943590	1.46	0.885604	1.80	0.931384
1.14	0.936416	1.48	0.885747	1.82	0.936845
1.16	0.929803	1.50	0.886227	1.84	0.942612
1.18	0.923728	1.52	0.887039	1.86	0.948687
1.20	0.918169	1.54	0.888178	1.88	0.955071
1.22	0.913106	1.56	0.889639	1.90	0.961766
1.24	0.908521	1.58	0.891420	1.92	0.968774
1.26	0.904397	1.60	0.893515	1.94	0.976099
1.28	0.900718	1.62	0.895924	1.96	0.983743
1.30	0.897471	1.64	0.898642	1.98	0.991708
1.32	0.894640	1.66	0.901668	2.00	1.000000

مهم‌ترین مزیت توزیع احتمالی گاما این است که این تابع خاصیت افزایشی (additive) داشته و می‌توان متغیرهای دیگری را که به صورت احتمالی به تابع اصلی اضافه می‌شوند در نظر گرفت. این موضوع بخصوص در شرایطی که رودخانه اصلی دارای انشعابات متعدد بوده و هر چه به طرف پائین دست می‌رویم شاخه جدیدی به رودخانه اضافه می‌شود و در نهایت دبی سیل در نقطه پائین دست موردنظر است حائزاهمیت بوده و در این وضعیت می‌توان از تابع توزیع احتمالی گاما استفاده کرد.

(۴) توزیع تئوری پیرسون تیپ ۳

معادله اساسی توزیع چگالی احتمالی پیرسون تیپ ۳ که برای اولین بار توسط فوستر (Foster) این توزیع را برای محاسبه حداکثر سیل لحظه‌ای بکار برد بصورت زیر می‌باشد:

$$f(x) = \{\frac{1}{\Gamma(b)}\}\,[\lambda^{b}(x-c)^{b-1}\,e^{-\lambda(x-c)}] \qquad (۵۵-۱۷)$$

در این معادله:

$$b = (\frac{2}{C_s})^2 \qquad (۵۶-۱۷)$$

$$C = \frac{\sigma}{\sqrt{b}} \qquad (۵۷-۱۷)$$

σ= انحراف از معیار داده‌ها

C_s = ضریب چولگی داده‌ها که توصیه می‌شود پس از محاسبه با توجه به تعداد داده‌ها (n) در ضریب $(1 + \frac{8.5}{n})$ ضرب شود.

ملاحظه می‌شود که اگر در توزیع تئوری پیرسون تیپ ۳ مقدار $C = 0$ باشد این تابع به تابع توزیع گاما تبدیل می‌شود.

توزیع تئوری پیرسون تیپ ۳ را نیز می‌توان با لگاریتم گرفتن از داده مشابه آن چه در حدود توزیع تئوری لوگ - نرمال گفته شد تغییر داده و به توزیع لوگ - نرمال پیرسون تیپ ۳ تبدیل کرد.

(۵) توزیع تئوری گامبل

در توزیع احتمالی گامبل، احتمال تجمعی وقوع یک پدیده $F(x)$ عبارت است از:

$$F(x) = e^{-e^{-\frac{(a+x)}{c}}} \qquad (۵۸-۱۷)$$

که اگر $\frac{a+x}{c}$ را برابر y در نظر بگیریم:

$$F(x) = e^{-e^{-y}} \qquad (۵۹-۱۷)$$

مقادیر c و a اعداد ثابتی هستند که به صورت زیر توصیف می‌شوند.

$$C = (\sqrt{6}/\pi)\sigma = 0.78009\ \sigma \qquad (۶۰-۱۷)$$

$$a = \gamma_c - x_{av} \qquad (۶۱-۱۷)$$

$$\gamma = 0.57721 \qquad (۶۲-۱۷)$$

در این معادلات σ انحراف از معیار و x_{av} میانگین داده‌هاست.

$$\gamma_c = (0.57721)\ (0.78009\ \sigma) = 0.45\ \sigma \qquad (۶۳-۱۷)$$

$$a = \gamma_c - x_{av} = 0.45\ \sigma - x_{av} \qquad (۶۴-۱۷)$$

لذا خواهیم داشت:

$$y = \frac{a + x}{c} = -\ln\ [-\ln\ F(x)] \qquad (۶۵-۱۷)$$

و اگر به جای احتمال عکس آن را برحسب دوره برگشت قرار دهیم:

$$y = -\ln\ \{\ln \frac{T}{T-1}\} \qquad (۶۶-۱۷)$$

بطور کلی کافی است که داده‌های هیدرولوژی را با یکی از توابع توزیع تئوری برازش بدهیم و اگر نکوئی برازش ثابت شد با حل تئوری معادله توزیع مقدار متغیر را به ازاء احتمالات مختلف بدست آوریم. امروزه نرم افزارهای کامپیوتری در اختیار است که علاوه بر برازش و ارائه نمایه‌های نکوئی برازش مقادیر متغیر را به ازاء آن توزیع تئوری در احتمالات مختلف محاسبه می‌کند. حل توابع تئوری توزیع‌های احتمال نیز به دلیل نامعین بودن محدوده انتگرال ساده نبوده و معمولاً به روش‌های عددی یا با استفاده از جداولی که برای این منظور تهیه شده است صورت می‌گیرد.

برای برازش داده‌های هیدرولوژی با توزیع‌های تئوری روش‌های گوناگونی وجود دارد که چهار روش آنها در هیدرولوژی کاربرد دارند. این روش‌ها عبارتند از:

- روش استفاده از پارامترهای توزیع
- روش ضرایب فراوانی
- روش گرافیک
- روش حداقل مربعات

اساسی‌ترین روش آن است که از ***پارامترهای توزیع*** استفاده کنیم. برای این کار بر اساس روش‌های احتمالات دو راه حل ***گشتاور*** (moment) و ***حداکثر درستنمائی*** (maximum likelihood) بکار برده می‌شود. روش گشتاور ساده است اما نتایج حاصله از آن بخصوص اگر تعداد داده‌ها کم باشد از دقت کمتری برخودار است.

برعکس روش حداکثر درستنمائی از دقت زیادی برخوردار است اما محاسبات آن بسیار پیچیده و وقت گیر است. بنابراین در هیدرولوژی که سرعت عمل ارجح است روش گشتاور مورد استفاده قرار می‌گیرد. شرح تئوری این روش‌ها از محدوده این کتاب خارج می‌باشد.

روش دوم برای برازش داده‌ها استفاده از ***ضرایب فراوانی*** (frequency factors) است. ضرایب فراوانی (K) بستگی به پارامترهای توزیع داشته و معمولاً از معادله زیر پیروی می‌کنند.

$$K = \frac{(x - \bar{x})}{S} \qquad (۶۷-۱۷)$$

در این معادله که مشابه با تبدیل Z در معادله ۱۷-۴۸ است $\bar{x}$ و S به ترتیب میانگین و انحراف از معیار داده‌ها می‌باشند. با کمی تغییر در معادله فوق خواهیم داشت:

$$x = \bar{x} + K.S \qquad (۶۸-۱۷)$$

بنابراین اگر میانگین و انحراف از معیار داده‌ها را داشته باشیم به ازاء مقادیر مختلف K که خود بستگی به احتمال وقوع داشته و از جداول ارائه شده قابل استخراج است می‌توانیم x را در احتمالات مختلف محاسبه کنیم. مثلاً اگر در یک توزیع خاص ضریب فراوانی به ازاء احتمال مورد نظر ۰/۸۶ و مقادیر میانگین و انحراف از معیار داده‌های بارندگی به ترتیب ۱۸۵ و ۶۹ میلی متر باشد مقدار x به ازاء آن احتمال ۲۴۴/۳ میلی متر محاسبه می‌شود.

روش سوم برای برازش داده‌ها با توزیع‌های تئوری ***روش نموداری یا گرافیکی*** است. برای این منظور کافی است که کاغذهای گراف مخصوصی را که محورهای آن بر اساس تویع‌های مختلف درجه بندی شده است انتخاب کرده و موقعیت داده‌های هیدرولوژی را از نظر مقدار و احتمال وقوع با استفاده از فرمول‌های تجربی (فرمول‌های ۱۷-۳۱ الی ۱۷-۳۷) روی این گراف‌ها مشخص نمائیم. اگر نقاط بدست آمده در امتداد یک خط مستقیم قرار گرفتند مشخص می‌شود که داده‌ها با تابع مورد نظر برازش دارند. برای برخی توزیع‌ها مانند توزیع نرمال، توزیع دوجمله‌ای و توزیع حدنهائی (گامبل) کاغذهای احتمال در دسترس می‌باشد.

در **روش چهارم** یا روش ***حداقل مربعات*** مجذور ابتدا داده‌ها به یک تابع توزیع برازش داده می‌شود سپس اختلاف بین ارقام حاصل از تابع و داده‌های مشاهده‌ای محاسبه و تابعی که این اختلاف در آن حداقل باشد انتخاب می‌گردد. ممکن است داده‌ها با چند توزیع احتمال برازش داشته باشند اما این که کدام توزیع را انتخاب کنیم خود مقوله جداگانه‌ای است که باید ***آزمون‌های نکوئی برازش*** (goodness-of-fit test) صورت گیرد. برای این منظور دو آزمون عمده وجود دارد که عبارتند از:

- آزمون کای مربع (Chi square) یا X^2

- آزمون کلموگروف - اسمیرنوف (Kolmogorov-Smirnow) یا K-S

از بین این دو آزمون K-S ساده‌تر می‌باشد. برای آزمون نکوئی برازش به روش K-S قدم اول آن است که از **فرض صفر** (null hypothesis) استفاده شود.

فرض صفر آن است که بین توزیع تئوری و توزیع داده‌های تجربی هیچ گونه اختلاف معنی داری وجود ندارد. قدم بعدی انتخاب یک سطح اعتماد (α) برای آزمون می‌باشد. مثلاً سطح اعتماد ۵ درصد ($\alpha = 0.05$) به این معنی است که آزمون در ۵ درصد موارد اشتباه و در ۹۵ درصد موارد صحیح است. بعد از انجام فرض صفر با توجه به موارد فوق برای هر یک از مشاهدات (داده‌ها) مقدار تخمین زده شده از توزیع را بدست آورده و حداکثر تفاوت (D_{max}) را بدون توجه به علامت آن با مقدار بحرانی آن (D) که از جدول ۱۷-۱۲ بدست می‌آید مقایسه می‌کنیم.

مقدار بحرانی D بستگی به تعداد نمونه‌ها (n) و سطح اعتماد (α) دارد. اگر $D_{max}<D$ باشد فرض صفر صحیح بوده و بین داده‌ها و توزیع تئوری تفاوتی وجود ندارد ولی اگر $D_{max}>D$ باشد فرض صفر صحیح نبوده و مثلاً با ۹۵ درصد اطمینان می‌توان گفت که داده‌ها از تابع تئوری مورد نظر تبعیت ندارند. این موضوع در مثال ۱۷-۱۲ شرح داده شده است.

جدول ۱۷-۱۲ مقدار بحرانی D برای آزمون کلموگروف - اسمیرنوف

تعداد نمونه (n)	$\alpha =$ 0.15	0.10	0.05	0.01
4	0.319	0.352	0.381	0.417
5	0.299	0.315	0.337	0.405
6	0.277	0.294	0.319	0.364
7	0.258	0.276	0.300	0.348
8	0.244	0.261	0.285	0.331
9	0.233	0.249	0.271	0.311
10	0.224	0.239	0.258	0.294
11	0.217	0.230	0.249	0.284
12	0.212	0.223	0.242	0.275
13	0.202	0.214	0.234	0.268
14	0.194	0.207	0.227	0.261
15	0.187	0.201	0.220	0.257
16	0.182	0.195	0.213	0.250
17	0.177	0.189	0.206	0.245
18	0.173	0.184	0.200	0.239
19	0.169	0.179	0.195	0.235
20	0.166	0.174	0.190	0.231
25	0.153	0.165	0.180	0.203
30	0.136	0.144	0.161	0.187
over 30	$0.768/\sqrt{n}$	$0.805/\sqrt{n}$	$0.886/\sqrt{n}$	$1.031/\sqrt{n}$

● مثال ۱۷-۱۲

داده‌های جدول ۱۷-۱۳ از بارندگی یک حوضه در اختیار است اولاً با چه احتمالی مقدار بارندگی ۴۵ سانتی متر یا بیشتر خواهد بود ثانیاً برازش دو توزیع نرمال و گاما را مقایسه کنید.

حل

الف - داده‌ها را به ترتیب نزولی ردیف می‌کنیم. (ستون دوم جدول ۱۷-۱۳)

ب - به هر کدام شماره ردیف (m) داده شود (ستون سوم جدول ۱۷-۱۳)

ج - از روی فرمول تجربی ویبول (فرمول ۱۷-۳۲) احتمال تجربی وقوع را بدست می‌آوریم (ستون ۴ جدول ۱۷-۱۳)

د - نقاط ستون‌های ۲ و ۴ را نسبت به روی یک کاغذ احتمال نرمال مشخص می‌کنیم.

ه - چون نقاط در امتداد یک خط مستقیم قرار می‌گیرند از بین آنها خط مذکور را می‌گذرانیم.

و - به لحاظ ظاهر داده‌ها از توزیع نرمال تبعیت می‌کنند.

ز - برای ۴۵ سانتی‌متر بارندگی احتمال وقوع ۲۸٪ یا دوره برگشت ۳/۶ سال می‌باشد.

برای مقایسه درجه برازش دو توزیع نرمال و گاما جدول ۱۷-۱۴ را تشکیل می‌دهیم. ستون‌های این جدول به شرح زیر است:

الف - ستون سوم ارقام مربوط به بارندگی است (PPT) که بصورت نزولی ردیف شده‌اند. ستون دوم شماره هر ردیف و ستون اول سال‌های وقوع می‌باشد.

ب - ستون چهارم احتمال تجربی است که برای هر کدام از داده‌ها از فرمول ویبول محاسبه شده است.

ج - ستون پنجم احتمال مربوط به هر یک از مقادیر بارندگی است که از توزیع تئوری نرمال محاسبه شده‌اند.

برای این منظور ابتدا میانگین $\overline{x}=41.28$ و انحراف از معیار داده‌ها $(S = 6.49)$ محاسبه شده‌اند و سپس با استفاده از فرمولی که در معادله ۱۷-۴۳ و شکل ۱۷-۱۶ برای توزیع نرمال داده شده است مقادیر احتمال تئوری به ازاء هر یک از مقادیر بارندگی ستون سوم محاسبه شده‌اند.

د - ارقام ستون ششم تفاوت بین احتمال تجربی (ستون ۴) و احتمال تئوری نرمال (ستون۵) می‌باشد (D).

ه - ارقام ستون هفتم احتمال تئوری توزیع گاما با توجه به فرمول‌های این روش می‌باشد. برای استفاده از این توزیع پارامترهای λ و β مورد نیاز بوده است که مقادیر آنها عبارت خواهند بود از:

$$\lambda = \frac{\overline{x}}{S^2} = 0.98 \quad \text{و} \quad \beta = \frac{\overline{x}^2}{S^2} = 40.45$$

و - ارقام ستون ۸ تفاوت احتمال تجربی و احتمال محاسبه شده با توزیع گاما می‌باشد.

حداکثر اختلاف (D) در توزیع گاما به مقدار ۰/۰۷۵ و مربوط به ردیف ۱۲ می‌باشد

جدول ۱۷-۱۳ تحلیل فراوانی داده‌های بارندگی

سال	بارندگی ردیف‌شده (inches)	ردیف (m)	فراوانی نمونه‌ها $F_s = [m/(n+1)]$
1983	54.41	1	0.026
1979	52.79	2	0.051
1975	52.13	3	0.077
1972	49.63	4	0.103
1977	49.42	5	0.128
1953	48.13	6	0.154
1958	47.87	7	0.179
1971	47.79	8	0.205
1973	46.06	9	0.231
1978	45.95	10	0.256
1956	45.90	11	0.282
1952	45.84	12	0.308
1967	44.82	13	0.333
1984	43.66	14	0.359
1969	43.36	15	0.385
1962	42.63	16	0.410
1951	42.06	17	0.436
1960	41.15	18	0.462
1961	41.05	19	0.487
1950	40.47	20	0.513
1982	40.43	21	0.538
1986	40.42	22	0.564
1966	40.00	23	0.590
1970	39.14	24	0.615
1980	38.80	25	0.641
1959	38.37	26	0.667
1981	37.83	27	0.692
1974	37.78	28	0.718
1968	35.45	29	0.744
1985	35.20	30	0.769
1963	34.95	31	0.795
1954	34.04	32	0.821
1987	33.40	33	0.846
1976	33.27	34	0.872
1955	33.03	35	0.897
1957	32.20	36	0.923
1964	29.88	37	0.949
1965	29.34	38	0.974

($D_{max} = 0.075$). حداکثر اختلاف در توزیع نرمال ۰/۰۷۱ و مربوط به ردیف ۲۹ می‌باشد. در جدول ۱۷-۱۲ به ازاء $n = 38$ و $\alpha = 0.05$ مقدار بحرانی D برابر ۰/۱۴ بدست می‌آید.

چون D_{max} در هر دو توزیع نرمال و گاما کمتر از ۰/۱۴ است لذا فرض صفر در مورد هر دو توزیع قابل قبول می‌باشد. به عبارت دیگر داده‌ها هم از توزیع نرمال و هم از توزیع گاما تبعیت می‌کنند. ولی اگر متوسط تفاوت‌ها را (D) در دو توزیع محاسبه کنیم مشاهده خواهد شد که متوسط D برای توزیع نرمال ۰/۰۲۳ و برای توزیع گاما ۰/۰۲۴ است و لذا نکوئی برازش داده‌ها با توزیع نرمال کمی بهتر از توزیع گاما می‌باشد.

جدول ۱۷-۱۴ آزمون نکوئی برازش کلموگروف - اسمیرنوف برای داده‌های بارندگی با دو توزیع نرمال و گاما.

سال	ردیف	PPT بارندگی	فراوانی	توزیع نرمال	D	توزیع گاما	D
1983	1	54.41	0.026	0.022	0.004	0.029	0.004
1979	2	52.79	0.051	0.038	0.013	0.046	0.005
1975	3	52.13	0.077	0.048	0.029	0.056	0.021
1972	4	49.63	0.103	0.099	0.003	0.104	0.001
1977	5	49.42	0.128	0.105	0.023	0.109	0.019
1953	6	48.13	0.154	0.146	0.008	0.146	0.008
1958	7	47.87	0.179	0.155	0.024	0.155	0.025
1971	8	47.79	0.205	0.158	0.047	0.157	0.048
1973	9	46.06	0.231	0.231	0.000	0.223	0.008
1978	10	45.95	0.256	0.236	0.020	0.228	0.028
1956	11	45.90	0.282	0.238	0.044	0.230	0.052
1952	12	45.84	0.308	0.241	0.066	0.233	0.075
1967	13	44.82	0.333	0.293	0.041	0.280	0.053
1984	14	43.66	0.359	0.357	0.002	0.340	0.019
1969	15	43.36	0.385	0.374	0.010	0.357	0.028
1962	16	42.63	0.410	0.417	0.007	0.398	0.012
1951	17	42.06	0.436	0.452	0.016	0.432	0.004
1960	18	41.15	0.462	0.508	0.046	0.487	0.025
1961	19	41.05	0.487	0.514	0.027	0.493	0.006
1950	20	40.47	0.513	0.549	0.036	0.529	0.016
1982	21	40.43	0.538	0.552	0.013	0.531	0.007
1986	22	40.42	0.564	0.552	0.012	0.532	0.032
1966	23	40.00	0.590	0.578	0.012	0.558	0.032
1970	24	39.14	0.615	0.629	0.013	0.611	0.005
1980	25	38.80	0.641	0.648	0.007	0.631	0.010
1959	26	38.37	0.667	0.672	0.006	0.657	0.009
1981	27	37.83	0.692	0.712	0.020	0.700	0.008
1974	28	37.78	0.718	0.705	0.013	0.692	0.026
1968	29	35.45	0.744	0.815	0.071	0.813	0.069
1985	30	35.20	0.769	0.825	0.056	0.824	0.055
1963	31	34.95	0.795	0.835	0.040	0.835	0.040
1954	32	34.04	0.821	0.867	0.047	0.871	0.051
1987	33	33.40	0.846	0.887	0.041	0.893	0.047
1976	34	33.27	0.872	0.891	0.019	0.898	0.026
1955	35	33.03	0.897	0.898	0.000	0.905	0.008
1957	36	32.20	0.923	0.919	0.004	0.928	0.005
1964	37	29.88	0.949	0.960	0.011	0.971	0.022
1965	38	29.34	0.974	0.967	0.008	0.977	0.003

۱۷-۹ تحلیل فراوانی وقایع

هدف از تحلیل فراوانی وقایع در هیدرولوژی بدست آوردن احتمال وقوع یک حادثه مانند حداکثر لحظه‌ای سیلاب، حداقل دبی رودخانه، حداکثر بارش ۲۴ ساعته و امثال آن است. برای این کار معمولاً از داده‌های سری‌های حداکثر (اکستریم) و یا سری‌های جزئی استفاده شده و آنها

را با توزیع‌های احتمالاتی نظری برازش می‌دهند. برای روشن شدن مطلب با ذکر یک مثال فرض می‌کنیم داده‌های حداکثر سیلاب لحظه‌ای سالانه (سری حداکثر) یک رودخانه را از سال ۱۹۳۵ لغایت ۱۹۷۳ بمدت ۳۸ سال (جدول ۱۷-۱۵) در اختیار داشته و می‌خواهیم برازش این داده‌ها را با ۵ توزیع تئوری زیر:

- تابع توزیع نرمال
- تابع توزیع لوگ - نرمال
- تابع توزیع گامبل تیپ ۱ (حد نهائی ۱)
- تابع توزیع پیرسون تیپ ۳
- تابع توزیع لوگ - پیرسون تیپ ۳

بررسی کرده و در نهایت مقدار سیل لحظه‌ای این رودخانه را بـه ازاء احـتمالات یـا دوره‌هـای بازگشت (۵ ، ۵۰ و ۱۰۰ سال) بدست آوریم. برای این منظور، ترکیبی از روش‌های گرافـیک، پارامتریک و ضرایب فراوانی را بکار می‌بریم.

از بین توزیع‌های فوق سه تای آنها دارای کاغذ احتمال بوده و لذا استفاده از روش گرافیکی در آنها ساده است. روش ضرایب فراوانی در مورد تمام این ۵ توزیع کاربرد دارد و روش استفاده از پارامترهای توزیع در تابع گامبل تیپ ۱ سهل‌تر خواهد بود. انتخاب نوع تابع بر اساس مقایسه نکوئی برازش خواهد بود.

جدول ۱۷-۱۵ حداکثر سیل لحظه‌ای مشاهده شـده در یک رودخـانه از سـال ۱۹۳۵ لغایت ۱۹۷۳ (ارقام برحسب فوت مکعب بر ثانیه می‌باشند)

سال	Q	سال	Q	سال	Q
1935	25100	1948	14000	1961	22800
1936	32400	1949	6600	1962	4040
1937	16300	1950	5080	1963	858
1938	24800	1951	5350	1964	13800
1939	903	1952	11000	1965	26700
1940	34500	1953	14300	1966	5480
1941	30000	1954	24200	1967	1900
1942	18600	1955	12400	1968	21800
1943	7800	1956	5660	1969	20700
1944	37500	1957	155000	1970	11200
1945	10300	1958	21800	1971	9640
1946	8000	1959	3080	1972	4790
1947	21000	1960	71500	1973	18100

برای انجام تحلیل فراوانی ابتدا داده‌ها را بصورت نزولی ردیف کرده و مقدار احتمال وقوع هر ردیف (F_s) را از فرمول ویبول محاسبه می‌کنیم. هر چند برخی از هیدرولوژیست‌ها عقیده دارند برای توزیع گامبل استفاده از فرمول گرینگورتن مناسب‌تر است. در جدول ۱۷-۱۶ ستون ۲ شماره ردیف، ستون ۳ احتمال تجربی محاسبه شده بر اساس فرمول ویبول و ستون ۴ دوره بازگشت (سال) و ستون ۵ مقادیر دبی می‌باشد که به ترتیب نزولی ردیف شده است. ارقام ستون ۶ لگاریتم مقادیر دبی است. بالاترین مقدار دبی مربوط به سال ۱۹۵۷ است که برابر ۱۵۵۰۰۰ می‌باشد. این دبی بسیار زیاد است بطوری که ممکن است لازم باشد آن را بعنوان داده *خارج از* **محدوده** (outlier) حذف کرد. داده‌های خارج از محدوده داده هائی هستند که بنظر می‌رسد هماهنگ با دیگر داده‌ها نبوده و از تابع توزیع تبعیت نمی‌کنند. یکی از مشکلات تحلیل فراوانی تشخیص داده‌های خارج از محدوده که خود روش‌های پیچیده آماری در پیش دارند لذا در اینجا ما تمام داده‌ها را لحاظ کرده و داده‌ای را حذف نخواهیم کرد.

جدول ۱۷-۱۶

سال	ردیف	F_s	T سال	دبی (cfs)	Log_{10} دبی
1957				155000	5.1903
1960	1	0.026	39.00	71500	4.8543
1944	2	0.051	19.50	37500	4.5740
1940	3	0.077	13.00	34500	4.5378
1936	4	0.103	9.75	32400	4.5105
1941	5	0.128	7.80	30000	4.4771
1965	6	0.154	6.50	26700	4.4265
1935	7	0.179	5.57	25100	4.3997
1938	8	0.205	4.88	24800	4.3945
⋮	⋮	⋮	⋮	⋮	⋮
1966	30	0.769	1.30	5480	3.7388
1951	31	0.795	1.26	5350	3.7284
1950	32	0.821	1.22	5080	3.7059
1972	33	0.846	1.18	4790	3.6803
1962	34	0.872	1.15	4040	3.6064
1959	35	0.897	1.11	3080	3.4886
1967	36	0.923	1.08	1900	3.2788
1939	37	0.949	1.05	903	2.9557
1963	38	0.974	1.03	858	2.9335
میانگین				16,421	4.0547
انحراف معیار				13,521	0.4287
چولگی				0.823	0.1057
ضریب تغییرات				1.934	−0.886
چولگی برای توزیع لوگ نرمال				3.028	

الف - توزیع نرمال در شکل ۱۷-۱۷ داده‌های احتمال تجربی (ارقام ستون ۳ جدول ۱۷-۱۶) نسبت به مقادیر دبی (ارقام ستون ۵) روی کاغذ احتمالات نرمال آورده شده‌اند علاوه بر این توزیع تئوری نرمال با میانگین ۱۶۴۲۱ و انحراف از معیار ۱۳۵۲۱ بصورت یک خط رسم شده است. بطوری که مشاهده می‌شود برازش چندان خوب نیست زیرا تعداد زیادی از داده‌های تجربی زیر خط قرار می‌گیرند. این امر به دلیل آن است که توزیع نرمال برای حالتی خوب است که چولگی آن‌ها صفر باشد ولی این داده‌ها دارای چولگی مثبت بوده لذا توزیع تئوری بطرف بالا کشیده شده است.

توزیع نرمال را با استفاده از ضرایب فراوانی و فرمول $x = \bar{x}+K.S$ نیز می‌توان مورد استفاده قرار داد. ضرایب فراوانی در توزیع نرمال به ازاء مقادیر احتمالات و دوره‌های برگشت مختلف در جدول ۱۷-۱۷ آورده شده است. مثلاً در این جدول برای دوره‌های برگشت ۵ و ۵۰ و ۱۰۰ سال مقادیر ضریب فراوانی K به ترتیب ۰/۸۴۲، ۲/۰۵۴ و ۲/۳۲۶ می‌باشد. بنابراین مقادیر حداکثر دبی لحظه‌ای در دوره‌های ۵، ۵۰ و ۱۰۰ سال با استفاده از توزیع نرمال برابر خواهند بود با:

$$Q_5=16421+0.842(13521)=27806$$

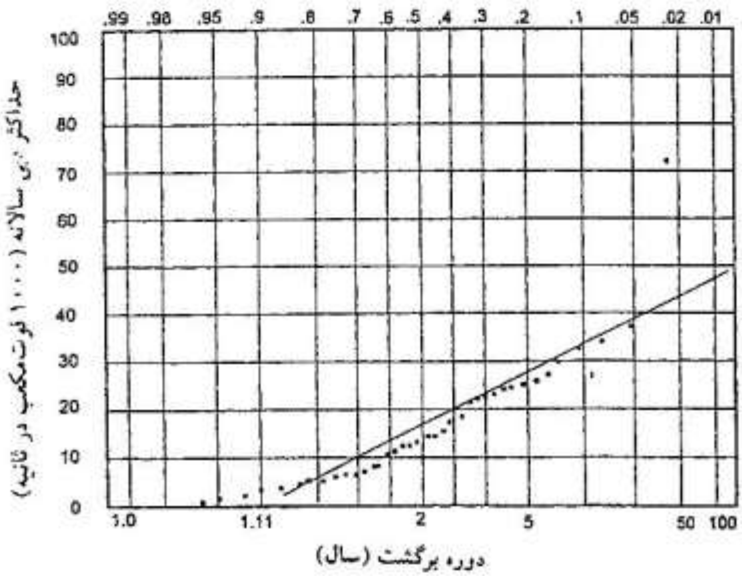
$$Q_{50}=16421+2.054(13521)=44193$$

$$Q_{100}=16421+2.326(13521)=47871$$

ب - توزیع لوگ نرمال توزیع لوگ - نرمال بر این فرض استوار است که چون داده‌ها دارای چولگی هستند اگر از داده‌های اولیه (x) لگاریتم گرفته شود ($y = \log x$) لگاریتم داده‌ها دارای توزیع نرمال خواهند بود. در شکل ۱۷-۱۸ داده‌های سیلاب (ستون ۳ و ۵ جدول ۱۷-۱۶) و خط مربوط به توزیع تئوری روی کاغذ لوگ - نرمال آورده شده است. راه دیگر آن است که با تبدیل داده‌ها به لگاریتم و مقادیر پارامترهای توزیع لوگ - نرمال ($\bar{y}$ و S_y) که به ترتیب ۴/۰۵۴۷ و ۰/۴۲۸۷ می‌باشند روی کاغذ نرمال به طریقی که قبلاً گفته شد لگاریتم داده‌ها را مورد بررسی قرار دهیم. در شکل ۱۷-۱۸ مشاهده می‌شود که داده‌ها از دوره بازگشت $T>1.6$ بخوبی با توزیع لوگ -نرمال مطابقت دارند ولی برای دوره‌های بازگشت $T<1.6$ این داده‌ها با توزیع لوگ - نرمال برازش ندارند.

توزیع لوگ - نرمال تا اندازه‌ای چولگی داده‌ها را برطرف می‌کند ولی بطور کامل آن را از بین نمی‌برد. برای محاسبه ضریب چولگی داده‌ها (C_s) در توزیع لوگ - نرمال باید از معادله زیر استفاده شود.

$$C_s=3(C_v)+(C_v)^3 \qquad (۱۷-۶۹)$$



شکل ۱۷-۱۷ برازش داده‌های تجربی سیلاب با توزیع تئوری نرمال

که در آن Cv ضریب تغییرات داده‌ها است. در این مثال ضریب تغییرات داده‌ها ۰/۸۲۳ و ضریب چولگی داده‌ها بر اساس فرمول ۱۷-۶۹ محاسبه شده و مقدار آن ۳/۰۲۸ بدست می‌آید.

با روش ضرایب فراوانی نیز می‌توان از توزیع لوگ - نرمال استفاده کرد. در جدول ۱۷-۱۸ مقادیر ضرایب فراوانی توزیع لوگ - نرمال به ازاء مقادیر چولگی (C_s)، و احتمالات یا دوره‌های برگشت مختلف آورده شده است. مثلاً برای داده‌های سیلاب که ضریب تغییرات و چولگی آن جهت استفاده در توزیع لوگ - نرمال به ترتیب ۰/۸۲۳ و ۳/۰۲۸ می‌باشد مقادیر ضرایب فراوانی در دوره‌های برگشت ۵ و ۱۰۰ سال ۰/۵۱ و ۳/۷۸ است. بنابراین حداکثر سیلاب تخمینی به روش ضرایب فراوانی در توزیع لوگ - نرمال برای دوره‌های برگشت ۵ و ۱۰۰ سال عبارتند از:

$$Q_5 = 16421 + 0.51(13521) = 23317$$

$$Q_{100} = 16421 + 3.78(13521) = 67530$$

بطوری که ملاحظه می‌شود. برای دوره ۵۰ ساله (Q_{50}) محاسبات به روش ضرایب فراوانی انجام نشده است زیرا بدست آوردن ضریب فراوانی با استفاده از روش درون یابی (interpolation) در روش لوگ - نرمال توصیه نمی‌شود ولی از توزیع تئوری لوگ - نرمال که در شکل ۱۷-۱۸ رسم شده است مقدار Q_{50} برابر ۵۴۵۰۰ بدست می‌آید.

جدول ۱۷-۱۷ ضرایب فراوانی در توزیع نرمال

احتمال (%)	دوره برگشت (سال)	ضریب فراوانی K	احتمال (%)	دوره برگشت (سال)	ضریب فراوانی K
0.01	10000	3.719	50.0	2.0	0.000
0.05	2000	3.291	55	1.82	-0.126
0.10	1000	3.090	60	1.67	-0.253
0.50	200	2.576	65	1.54	-0.385
1.0	100	2.326	70	1.43	-0.524
2.5	40	1.960	75	1.33	-0.674
5.0	20	1.645	80	1.25	-0.842
10.0	10	1.282	85	1.18	-1.036
15.0	6.67	1.036	90	1.11	-1.282
20.0	5	0.842	95	1.05	-1.645
25.0	4	0.674	97.5	1.03	-1.960
30.0	3.33	0.524	99.0	1.01	-2.326
35.0	2.86	0.385	99.5	1.005	-2.576
40.0	2.5	0.253	99.9	1.001	-3.090
45.0	2.22	0.126	99.95	1.0005	-3.291
50.0	2.0	0.000	99.99	1.0001	-3.719

ج - توزیع حد نهائی (گامبل تیپ ۱) توزیع حد نهائی تیپ یک (extreme value) که بطور خلاصه روش گامبل نامیده می‌شود برای داده‌هائی استفاده می‌شود که چولگی آنها مثبت است (برای داده‌های با چولگی منفی از توزیع حد نهائی تیپ ۳ استفاده می‌شود. مثلاً از توزیع گامبل تیپ ۳ برای تحلیل فراوانی خشکسالی استفاده می‌شود. روش گامبل تیپ ۱ با هر سه روش - گرافیکی، ضرایب فراوانی، و استفاده از پارامترهای توزیع قابل محاسبه است. یکی از دلایلی که روش گامبل را ساده می‌کند این است که فرض می‌شود چولگی داده‌ها ثابت و مقدار آن برابر $C_s = 1.1396$ است. برای توزیع گامبل تیپ ۱ نیز کاغذ احتمالات وجود دارد که در شکل ۱۷-۱۹ داده‌ها (ستون ۳ و ۵ جدول ۱۷-۱۶) روی آن آورده شده و خط تابع توزیع گامبل ۱ نیز رسم شده است. مشاهده می‌شود داده‌ها با این توزیع برازش خوبی دارند.

برای استفاده از توزیع گامبل تیپ ۱ نیز جدول ضرایب فراوانی وجود دارد که بر اساس تعداد داده‌ها و احتمال یا دوره برگشت مورد نظر می‌توان ضرایب فراوانی را بدست آورد. جدول ۱۷-۱۶ ضرایب فراوانی برای استفاده از روش گامبل تیپ ۱ را نشان می‌دهد. مثلاً به ازاء تعداد داده‌های این مثال (n=38) برای دوره‌های بازگشت ۵، ۵۰ و ۱۰۰ سال مقادیر K به ترتیب برابر ۰/۸۴۳ ، ۲/۹۵۷ و ۳/۵۷۲ می‌باشد (تعداد ۳۸ سال داده بین ارقام ۳۵ و ۴۰ سال درون یابی شده است). بنابراین مقادیر تخمین سیلاب به ازاء دوره‌های برگشت مورد نظر عبارتند از:

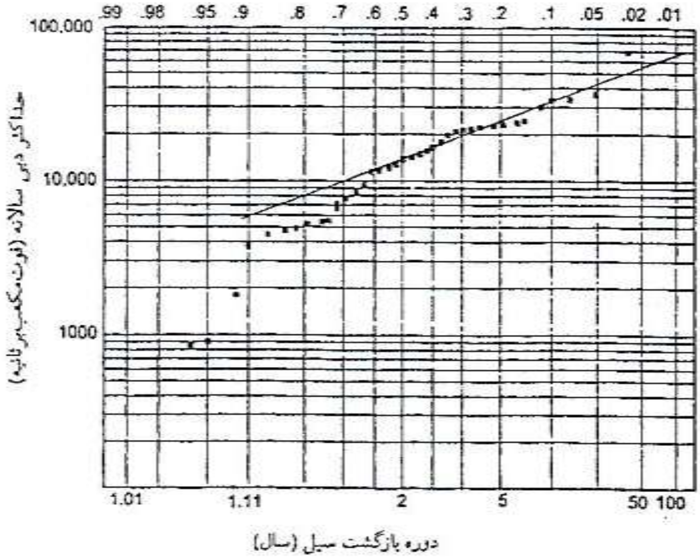
$$Q_5 = 16421 + 0.843(13521) = 27819$$

$$Q_{50} = 16421 + 2.957(13521) = 56403$$

$$Q_{100} = 16421 + 3.572(13521) = 64718$$

جدول ۱۷-۱۸ ضرایب فراوانی (K) برای توزیع لوگ - نرمال

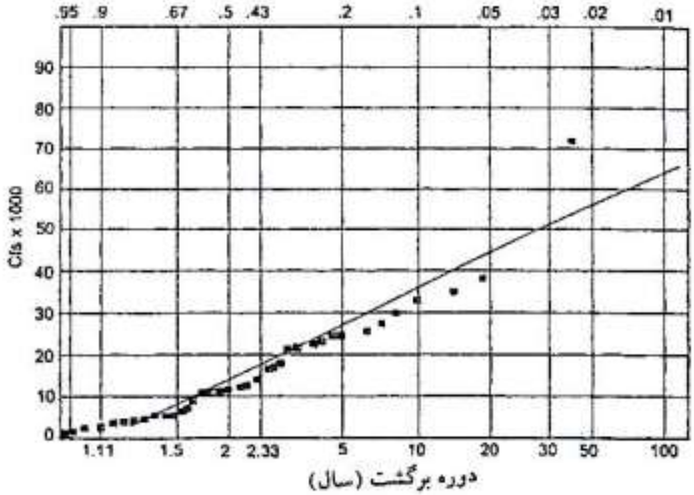
C_s	CV	دوره برگشت (سال)						
		1.01	1.05	1.25	2	5	20	100
		احتمال بیش از ...						
		0.99	0.95	0.80	0.50	0.20	0.05	0.01
		−	−	−	−	+	+	+
0.0	0.000	2.33	1.65	0.84	0.00	0.84	1.64	2.33
0.1	0.033	2.25	1.62	0.85	0.02	0.84	1.67	2.40
0.2	0.067	2.18	1.59	0.85	0.03	0.83	1.70	2.48
0.3	0.100	2.11	1.56	0.85	0.05	0.82	1.72	2.55
0.4	0.136	2.04	1.52	0.85	0.07	0.81	1.75	2.63
0.5	0.166	1.97	1.49	0.86	0.08	0.80	1.77	2.70
0.6	0.197	1.91	1.46	0.85	0.10	0.79	1.79	2.77
0.7	0.230	1.85	1.43	0.85	0.11	0.78	1.81	2.84
0.8	0.262	1.79	1.40	0.84	0.13	0.77	1.82	2.90
0.9	0.292	1.74	1.37	0.84	0.14	0.76	1.84	2.97
1.0	0.324	1.68	1.34	0.83	0.15	0.74	1.85	3.04
1.1	0.351	1.63	1.31	0.83	0.16	0.73	1.86	3.09
1.2	0.381	1.58	1.29	0.82	0.17	0.72	1.87	3.15
1.3	0.409	1.54	1.26	0.82	0.18	0.71	1.88	3.21
1.4	0.436	1.50	1.23	0.81	0.19	0.69	1.88	3.26
1.5	0.462	1.46	1.21	0.81	0.20	0.68	1.89	3.30
1.6	0.490	1.42	1.19	0.80	0.21	0.67	1.89	3.35
1.7	0.517	1.38	1.16	0.79	0.22	0.65	1.89	3.40
1.8	0.544	1.34	1.14	0.78	0.22	0.64	1.89	3.44
1.9	0.570	1.31	1.12	0.78	0.23	0.63	1.89	3.48
2.0	0.596	1.28	1.10	0.77	0.24	0.61	1.89	3.52
2.1	0.620	1.25	1.08	0.76	0.24	0.60	1.89	3.55
2.2	0.643	1.22	1.06	0.76	0.25	0.59	1.89	3.59
2.3	0.667	1.20	1.04	0.75	0.25	0.58	1.88	3.62
2.4	0.691	1.17	1.02	0.74	0.26	0.57	1.88	3.65
2.5	0.713	1.15	1.00	0.74	0.26	0.56	1.88	3.67
2.6	0.734	1.12	0.99	0.73	0.26	0.55	1.87	3.70
2.7	0.755	1.10	0.97	0.72	0.27	0.54	1.87	3.72
2.8	0.776	1.08	0.96	0.72	0.27	0.53	1.86	3.74
2.9	0.796	1.06	0.95	0.71	0.27	0.52	1.86	3.76
3.0	0.918	1.04	0.93	0.70	0.28	0.51	1.85	3.78
3.2	0.857	1.01	0.90	0.69	0.28	0.49	1.84	3.81
3.4	0.895	0.98	0.88	0.68	0.29	0.47	1.82	3.84
3.6	0.930	0.95	0.86	0.67	0.29	0.46	1.81	3.87
3.8	0.966	0.92	0.84	0.66	0.29	0.44	1.80	3.89
4.0	1.000	0.90	0.82	0.65	0.29	0.42	1.78	3.91
4.5	1.081	0.84	0.78	0.63	0.30	0.39	1.75	3.94
5.0	1.155	0.80	0.74	0.60	0.30	0.36	1.71	3.96



شکل ۱۷-۱۸ برازش داده‌های تجربی سیلاب با توزیع تئوری لوگ-نرمال.

جدول ۱۷-۱۹ ضرایب فراوانی برای استفاده در روش گامبل تیپ ۱

n	دوره برگشت(سال)						
	5	10	15	20	25	50	100
	احتمال بیش از ...						
	0.20	0.10	0.067	0.05	0.04	0.02	0.01
15	0.967	1.703	2.117	2.410	2.632	3.321	4.005
20	0.919	1.625	2.023	2.302	2.517	3.179	3.836
25	0.888	1.575	1.963	2.235	2.444	3.088	3.729
30	0.866	1.541	1.922	2.188	2.393	3.026	3.653
35	0.851	1.516	1.891	2.152	2.354	2.979	3.598
40	0.838	1.495	1.866	2.126	2.326	2.943	3.554
45	0.829	1.478	1.847	2.104	2.303	2.913	3.520
50	0.820	1.466	1.831	2.086	2.283	2.889	3.491
55	0.813	1.455	1.818	2.071	2.267	2.869	3.467
60	0.807	1.446	1.806	2.059	2.253	2.852	3.446
65	0.801	1.437	1.796	2.048	2.241	2.837	3.429
70	0.797	1.430	1.788	2.038	2.230	2.824	3.413
75	0.792	1.423	1.780	2.029	2.220	2.812	3.400
80	0.788	1.417	1.773	2.020	2.212	2.802	3.387
85	0.785	1.413	1.767	2.013	2.205	2.793	3.376
90	0.782	1.409	1.762	2.007	2.198	2.785	3.367
95	0.780	1.405	1.757	2.002	2.193	2.777	3.357
100	0.779	1.401	1.752	1.998	2.187	2.770	3.349
∞	0.719	1.305	1.635	1.866	2.044	2.592	3.137



شکل ۱۷-۱۹ برازش داده‌های تجربی سیلاب با توزیع تئوری گامبل تیپ ۱

توجه شود که در روش گامبل تیپ ۱ درون یابی بین دوره‌های برگشت توصیه نمی‌شود ولی به ازاء یک دوره برگشت مشخص درون یابی بین تعداد سالهای آماری قابل قبول است. مثلاً اگر می‌خواستیم سیلاب را به ازاء دوره برگشت ۷۵ سال بدست آوریم نمی‌توانستیم از جدول ضرایب فراوانی ۱۷-۱۹ استفاده نمائیم.

روش گامبل تیپ ۱ یکی از ساده‌ترین توزیع‌ها برای برازش داده‌ها با استفاده از پارامترهای توزیع است. فرمول هائی که برای این منظور استفاده می‌شود عبارتند از:

$$P = e^{-e^{-[(u+x)/\alpha]}} \qquad (۱۷-۷۰)$$

$$u = 0.5772\,\alpha - \overline{x} \qquad (۱۷-۷۱)$$

$$\alpha = 0.7797\,S \qquad (۱۷-۷۲)$$

که P احتمال وقوع، $\overline{x}$ میانگین داده‌ها، S انحراف از معیار داده‌ها و x مقدار متغیری است که برای آن احتمال وقوع (p) محاسبه می‌شود. مثلاً برای داده‌های مثال فوق اگر بخواهیم بدانیم دوره برگشت سیلاب‌های مربوط به ۴۰۰۰۰ چقدر است خواهیم داشت:

$\overline{x} = 16421$

$S = 13521$

$\alpha = 0.7797\,S = 0.7797(13521) = 10542.3$

$u = 0.5772\,(10542.3) - 16421 = -10335.9$

$x = 40000$

$$P = e^{-e^{-\left[\frac{-10336.9 + 40000}{10542.3}\right]}}$$

$$P = 0.9418$$

یعنی با احتمال ۹۴/۱۸ درصد دبی سالانه در این رودخانه به ۴۰۰۰۰ خواهد رسید و با احتمال $1 - 0.9418 = 0.058$ یا ۵/۸ درصد حداکثر دبی لحظه‌ای از ۴۰۰۰۰ بیشتر خواهد شد که دوره برگشت آن ۱۷/۲ سال است، زیرا:

$$T = \frac{1}{1 - P} = \frac{1}{1 - 0.9418} = \frac{1}{0.058} = 17.2$$

لذا:

$$Q_{17.2} = 40000$$

در روش گامبل تیپ ۱ میانگین داده‌ها بجای احتمال ۰/۵ منطبق بر احتمال ۰/۵۷ و یا دوره برگشت ۲/۳۳ سال است (بجای ۲ سال) بهمین دلیل در محاسبات تحلیل فراوانی نمایه $Q_{2.33}$ بجای Q_2 زیاد بکار برده می‌شود.

د - توزیع پیرسون تیپ ۳ توزیع پیرسون تیپ ۳ هم با داده‌های اولیه و هم با داده‌های تبدیل شده به لگاریتم مورد استفاده قرار می‌گیرد. در حالت اول توزیع را پیرسون تیپ ۳ و در حالت دوم توزیع لوگ-پیرسون تیپ ۳ نامند. برازش توزیع تیپ ۳ معمولاً با استفاده از ضرایب فراوانی انجام می‌شود زیرا برای این مورد کاغذ خاص پیرسون تیپ ۳ وجود ندارد. ضرایب فراوانی توزیع پیرسون تیپ ۳ بر اساس ضریب چولگی و احتمال وقوع (یا دوره برگشت) در جدول ۱۷-۲۰ آورده شده است. ضریب چولگی (C_s) در صورتی که تعداد داده‌ها کمتر از ۵۰ باشد، باید اصلاح گردد. این ضریب با استفاده از فرمول زیر اصلاح می‌شود. در این مثال ضریب اصلاح نشده چولگی ۱/۹۳۴ محاسبه می‌شود که اگر بخواهیم آن را برای $n = 38$ اصلاح کنیم مقدار آن ۲/۳۶ خواهد شد زیرا:

$$C'_s=(1+\frac{8.5}{n})1.934$$

$$C'_s=(1+\frac{8.5}{38})1.934=2.36$$

به ازاء مقدار چولگی اصلاح شده ۲/۳۶ ضرایب فراوانی K در دوره‌های برگشت ۵ و ۵۰ و ۱۰۰ سال با استفاده از روش درون یابی به ترتیب ۰/۵۴۴، ۳/۰۱ و ۳/۷۸۱ می‌باشد لذا:

$$Q_5=16421+0.544(13521)=23776$$

$$Q_{50}=16421+3.01(13521)=57119$$

$$Q_{100}=16421+3.781(13521)=67544$$

این دبی‌ها را می‌توان روی کاغذ لوگ - نرمال یا نیمه لگاریتمی رسم کرد ولی نتیجه حاصله یک

منحنی خواهد شد نه یک خط مستقیم.

هـ - توزیع لوگ - پیرسون تیپ ۳ روش لوگ - پیرسون تیپ ۳ مشابه با توزیع پیرسون تیپ ۳ می‌باشد با این تفاوت که در این جا با لگاریتم داده‌های اولیه کار می‌شود ولی مشکل اساسی در هنگام کار با لگاریتم داده‌ها تعیین ضریب چولگی است. برای این کار سه راه حل وجود دارد. یکی این که ضریب چولگی لگاریتم داده‌ها را بر اساس فرمول‌های ارائه شده محاسبه کنیم که در مورد این مثال ۰/۸۸۶ خواهد شد در این جا مسأله اصلاح آن به دلیل این که تعداد داده‌ها کمتر از ۵۰ می‌باشد وجود خواهد شد ولی فرمولی که برای اصلاح ضریب چولگی گفته شد در این جا صادق نمی‌باشد. راه حل دوم استفاده از ضرایب چولگی است که بر اساس تحلیل‌های منطقه‌ای برای هر منطقه محاسبه شده است مثلاً در امریکا برای نقاط مختلف آن کشور منحنی‌های ضریب چولگی رسم شده است ولی چنین کاری در ایران انجام نشده است. روش سوم آن است که ضریب چولگی داده‌های لگاریتم گرفته شده را حساب کنیم و عدد حاصله از تحلیل منطقه را نیز در نظر گرفته و میانگین آن‌ها را ملاک قرار دهیم. بنابراین ناچاراً در این مثال ما ضریب چولگی داده‌های ستون ۶ جدول ۱۷-۱۶ را که ۰/۸۸۶- است در نظر گرفته و مقادیر ضرایب K را برای دوره‌های برگشت ۵ و ۵۰ و ۱۰۰ سال از جدول ۱۷-۲۰ بدست می‌آوریم که به ترتیب برابر ۰/۸۵۷، ۱/۷۰، ۱/۸۶ می‌باشند. لذا خواهیم داشت:

$$y = \log x$$

$$\overline{y} = 4.0547$$

$$S_y = 0.4287$$

$$\log Q_5 = 4.0547+0.857(0.4287) = 4.4221$$

$$\log Q_{50} = 4.0547+1.70(0.4287) = 4.7835$$

$$\log Q_{100} = 4.0547+1.86(0.4287) = 4.8521$$

که اگر از آنها آنتی لگاریتم بگیریم مقادیر دبی بدست می‌آید.

$$Q_5 = 10^{4.4221} = 26430$$

$$Q_{50} = 10^{4.7835} = 60744$$

$$Q_{100} = 10^{4.8521} = 71135$$

یادآور می‌شود که برای توزیع لگاریتم پیرسون تیپ ۳ نیز کاغذ احتمال وجود نداشته ولی می‌توان نتایج را روی کاغذ لوگ - نرمال یا نیمه لگاریتمی رسم کرد.

جدول ۱۷-۲۰ ضرایب فراوانی برای استفاده از توزیع پیرسون تیپ ۳

چولگی C_s	Return period (years) دوره بازگشت (سال)						
	2	5	10	25	50	100	200
	Exceedence probability احتمال بیش از—						
	0.50	0.20	0.10	0.04	0.02	0.01	0.005
3.0	−0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.9	−0.390	0.440	1.195	2.277	3.134	4.013	4.909
2.8	−0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.7	−0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.6	−0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.5	−0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.4	−0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.3	−0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.2	−0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.1	−0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.0	−0.307	0.609	1.302	2.219	2.912	3.605	4.298
1.9	−0.294	0.627	1.310	2.207	2.881	3.553	4.223
1.8	−0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.7	−0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.6	−0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.5	−0.240	0.690	1.333	2.146	2.743	3.330	3.910
1.4	−0.225	0.705	1.337	2.128	2.706	3.271	3.828
1.3	−0.210	0.719	1.339	2.108	2.666	3.211	3.745
1.2	−0.195	0.732	1.340	2.087	2.626	3.149	3.661
1.1	−0.180	0.745	1.341	2.066	2.585	3.087	3.575
1.0	−0.164	0.758	1.340	2.043	2.542	3.022	3.489
0.9	−0.148	0.769	1.339	2.018	2.498	2.957	3.401
0.8	−0.132	0.780	1.336	1.993	2.453	2.891	3.312
0.7	−0.116	0.790	1.333	1.967	2.407	2.824	3.223
0.6	−0.099	0.800	1.328	1.939	2.359	2.755	3.132
0.5	−0.083	0.808	1.323	1.910	2.311	2.686	3.041
0.4	−0.066	0.816	1.317	1.880	2.261	2.615	2.949
0.3	−0.050	0.824	1.309	1.849	2.211	2.544	2.856
0.2	−0.033	0.830	1.301	1.818	2.159	2.472	2.763
0.1	−0.017	0.836	1.292	1.785	2.107	2.400	2.670
0.0	0.000	0.842	1.282	1.751	2.054	2.326	2.576
−0.1	0.017	0.846	1.270	1.716	2.000	2.252	2.482
−0.2	0.033	0.850	1.258	1.680	1.945	2.178	2.388
−0.3	0.050	0.853	1.245	1.643	1.890	2.104	2.294
−0.4	0.066	0.855	1.231	1.606	1.834	2.029	2.201
−0.5	0.083	0.856	1.216	1.567	1.777	1.955	2.108
−0.6	0.099	0.857	1.200	1.528	1.720	1.880	2.016
−0.7	0.116	0.857	1.183	1.488	1.663	1.806	1.926
−0.8	0.132	0.856	1.166	1.448	1.606	1.733	1.837
−0.9	0.148	0.854	1.147	1.407	1.549	1.660	1.749
−1.0	0.164	0.852	1.128	1.366	1.492	1.588	1.664
−1.1	0.180	0.848	1.107	1.324	1.435	1.518	1.581
−1.2	0.195	0.844	1.086	1.282	1.379	1.449	1.501
−1.3	0.210	0.838	1.064	1.240	1.324	1.383	1.424
−1.4	0.225	0.832	1.041	1.198	1.270	1.318	1.351
−1.5	0.240	0.825	1.018	1.157	1.217	1.256	1.282
−1.6	0.254	0.817	0.994	1.116	1.166	1.197	1.216
−1.7	0.268	0.808	0.970	1.075	1.116	1.140	1.155
−1.8	0.282	0.799	0.945	1.035	1.069	1.087	1.097
−1.9	0.292	0.788	0.920	0.996	1.023	1.037	1.044
−2.0	0.307	0.777	0.895	0.959	0.980	0.990	0.995
−2.1	0.319	0.765	0.869	0.923	0.939	0.946	0.949
−2.2	0.330	0.752	0.844	0.888	0.900	0.905	0.907
−2.3	0.341	0.739	0.819	0.855	0.864	0.867	0.869
−2.4	0.351	0.725	0.795	0.823	0.830	0.832	0.833
−2.5	0.360	0.711	0.771	0.793	0.798	0.799	0.800
−2.6	0.368	0.696	0.747	0.764	0.768	0.769	0.769
−2.7	0.376	0.681	0.724	0.738	0.740	0.740	0.741
−2.8	0.384	0.666	0.702	0.712	0.714	0.714	0.714
−2.9	0.390	0.651	0.681	0.683	0.689	0.690	0.690
−3.0	0.396	0.636	0.666	0.666	0.666	0.667	0.667

و - ساختن کاغذهای احتمالی برای هر یک از توزیع‌هایی که در آن رابطه ضریب فراوانی (K) و دوره برگشت (T) مشخص باشد (K-T relation) می‌توان کاغذ احتمال را ساخت. برای این منظور به ازاء مقادیر مختلف K مقادیر T وابسته به آن محاسبه می‌شود. سپس یک دستگاه محورهای متعامد را انتخاب کنید که محور افقی آن به مقادیر K (از صفر تا ۷+) اختصاص داده شده باشد. حال به ازاء T های مورد نظر (مثلاً ۲، ۵، ۱۰، ۵۰ و ... سال) مقدار K ر بدست آورده و روی محور علامت‌گذاری کنید. محور عمودی را نیز می‌توانید بصورت ساده و یا لگاریتمی مدرج کنید. چنین کاغذی که محور افقی آن بر حسب رابطه T و K برای دوره‌های برگشت مدرج شده است کاغذ احتمال برای آن توزیع بخصوص خواهد بود.

۱۰-۱۷ ریسک

واژه ریسک (risk) در هیدرولوژی به معنی احتمال وقوع یک حادثه در یک دوره زمانی مشخص است. مثلاً گفته می‌شود ریسک این که یک باران یا سیل ۱۰۰ ساله در ۵ سال آینده رخ دهد چقدر است. بر اساس آنچه تا بحال گفته شد برای واقعه‌ای که دوره بازگشت آن ۱۰۰ سال می‌باشد احتمال وقوع $P = 0.01$ است. یعنی شانس وقوع آن در هر سال یک درصد می‌باشد. اگر شانس وقوع حادثه‌ای P باشد شانس عدم وقوع آن $1 - P$ خواهد بود. لذا احتمال این که در ۲ سال آینده چنین واقعه‌ای رخ ندهد برابر خواهد بود با:

$$(1 - P)(1 - P) = (1 - P)^2 \qquad (۱۷-۷۳)$$

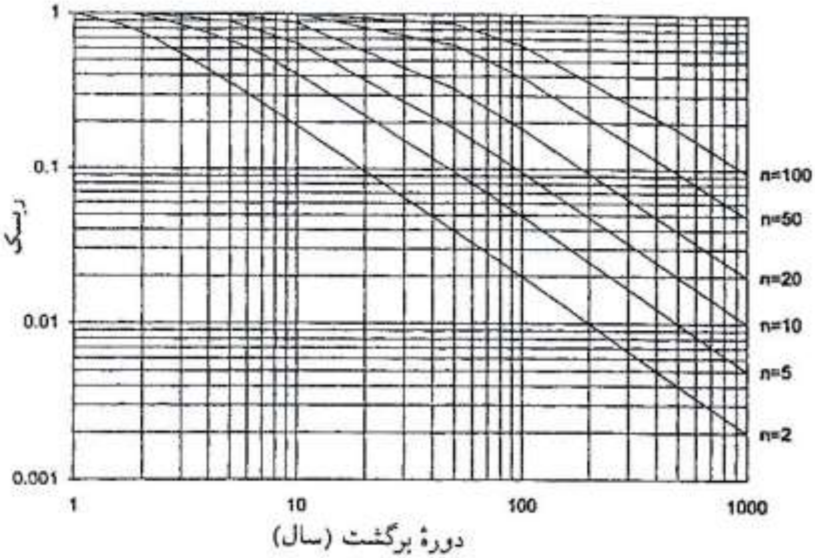
بهمین ترتیب احتمال این که واقعه مذکور در n سال آینده رخ ندهد $(1 - P)^n$ می‌باشد. بنابراین احتمال رخ دادن همین حادثه در طی n سال آینده برابر خواهد بود با:

$$J = 1 - (1 - P)^n \qquad (۱۷-۷۴)$$

این معادله بیان می‌دارد که اگر احتمال قوع حادثه‌ای در هر سال P باشد ریسک این که در n سال آینده حداقل یکبار چنین حادثه‌ای رخ دهد J خواهد بود و ریسک این که در n سال آینده حداقل k مرتبه حادثه مذکور رخ دهد عبارت است از:

$$J_k = \frac{n!}{k!\,(n - k)!} (1 - P)^{n-k} P^k \qquad (۱۷-۷۵)$$

در شکل ۱۷-۲۰ ریسک بعنوان تابعی از دوره بازگشت (T) و دوره زمانی (n) داده شده است. مثلاً مطابق این شکل برای حادثه‌ای که دوره بازگشت آن ۲۰۰ سال است ریسک وقوع در ۱۰ سال آینده ۰/۰۵ یا ۵ درصد است.



شکل ۱۷-۲۰ ریسک وقوع در n سال آینده بعنوان تابعی از دوره بازگشت (T)

● مثال ۱۷-۱۴

ریسک وقوع حداکثر لحظه‌ای سیل سالانه با دوره بازگشت ۲۰ سال در ۳ سال آینده چقدر است.

$P = 1/T = 1/20 = 0.05$

$n = 3$

$J = 1-(1-0.05)^3 = 14.3\%$

در مثال فوق محاسبه شد که ۱۴ درصد (یا کمی بیشتر) احتمال وجود دارد که سیلابی با دوره برگشت ۲۰ سال در طی ۳ سال آینده یکبار رخ دهد. بهمین ترتیب اگر احتمال این را که یک واقعه ۱۰۰ ساله ($p = 0.01$) در طی ۱۰۰ سال آینده ($n = 100$) حداقل یکبار رخ دهد محاسبه کنیم ریسک ۶۳ درصد خواهد بود که کمتر از متوسط آن یعنی یک بار در هر ۱۰۰ سال می‌باشد.

$J = 1 - (1 - 0.01)^{100} = 0.63$

حال این سوال مطرح است که آیا این ریسک قابل پذیرش است. تاسیساتی که در رابطه با آب احداث می‌شوند باید قادر باشند در مقابل سیلهایی با احتمال وقوع مشخص مقاومت نمایند. آگاهی از روابط مقدار - فراوانی کمک می‌کند تا در طراحی ساختمانهایی مانند سدها، جاده‌ها، پلها، آبگذرها و بندهای انحرافی ابعادی به کار گرفته شود که خطرات ناشی از سیل تا حد ممکن کاهش یابد.

روش معمول در طراحی سدها استفاده از حداکثر سیل ممکن یا در واقع نامحتملترین سیل

ممکن (PMF) می‌باشد. در این روش تأسیسات سد بر اساس بزرگترین سیلی که ممکن است در منطقه جاری شود بدون درنظر گرفتن درجه احتمال طراحی می‌شود. روش دیگر که در صورت وجود داده‌های مشاهده شده می‌توان از آن استفاده نمود روش به اصطلاح تحلیل فراوانی است که در آن برای هر یک از تاسیسات دوره برگشت مشخصی در نظر گرفته می‌شود و بر اساس تجزیه و تحلیل آماری داده‌ها مقدار سیل متناسب با آن دوره برگشت محاسبه می‌شود.

انتخاب دوره برگشت با توجه به تجزیه و تحلیلهای اقتصادی صورت می‌گیرد بطور کلی دوره برگشت برای تاسیسات مختلف به صورت زیر انتخاب می‌شود.

سدهای خاکی	۱۰۰۰ سال
سدهای بتونی و سنگ چین باملات	۵۰۰ سال
سدهای کوتاه	۱۰۰ سال
بندهای کوچک	۲۰ سال

اگر یک طرح بر اساس واقعه‌ای با دوره برگشت T ساله طراحی شود و عمر طرح n سال باشد احتمال این که واقعه مذکور حداقل یکبار در عمر پروژه رخ دهد از فرمول ۱۷-۳۱ قابل محاسبه است.

به عنوان مثال چنانچه سرریز سدی که عمر مفید آن ۱۰۰ سال پیش بینی می‌شود بر اساس سیل ۵۰۰ ساله طراحی شده باشد احتمال این که چنین سیلی در عمر ۱۰۰ ساله سد حداقل یکبار رخ دهد ۱۸ درصد خواهد بود زیرا:

$T = 500$ سال

$$P = 1 - \left(1 - \frac{1}{500}\right)^{100}$$

$$P = 0.18$$

این احتمال را خطر شکست یا ریسک شکست (failure) **گویند.** فرمول ۱۷-۳۱ برای وضعیتی است که طراحی بر مبنای سریهای ماکزیمم سالانه انجام شده باشد در صورتی که طراحی بر اساس سریهای جزئی صورت پذیرفته باشد فرمول بصورت زیر خواهد بود.

$$P = 1 - \left(1 - \frac{1}{T}\right)^{nk} \qquad (۱۷-۷۶)$$

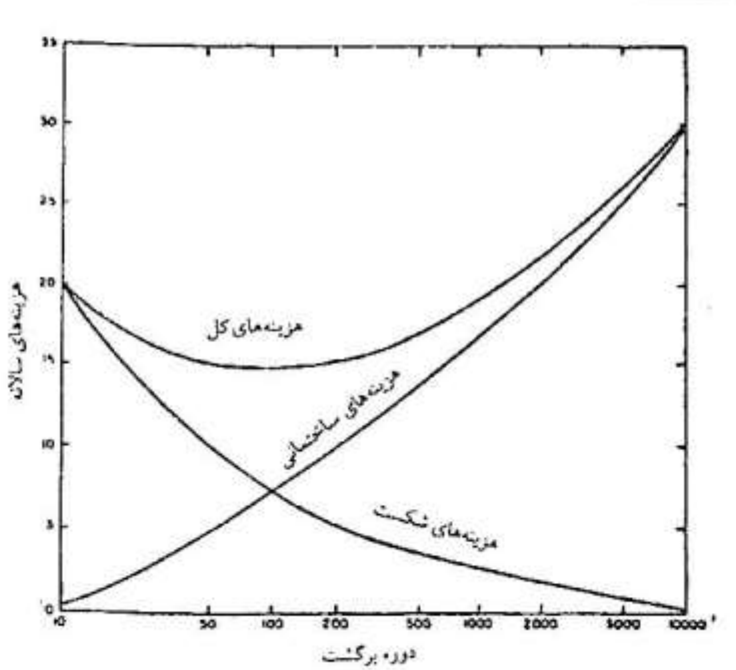
که k متوسط تعداد وقایع در هر سال است. در طرحهای هیدرولوژی ریسک مجاز برای شکست بستگی به عمر تاسیسات و دوره بازگشت دارد و مقدار آن از جدول ۱۷-۲۱ قابل محاسبه است.

برای مثال مطابق این جدول اگر بخواهیم ریسک مجاز برای شکست پروژه‌ای ۱۰ درصد و عمر پروژه ۲۵ سال باشد باید طراحی بر اساس دوره برگشت ۲۳۸ ساله صورت گیرد. البته انتخاب صحیح دوره برگشت در طراحی باید بر اساس تجزیه و تحلیلهای اقتصادی هزینه‌های سالانه صورت گیرد(شکل ۱۷-۲۱). مثلاً هر چه دوره برگشت بزرگتر انتخاب شود هزینه‌های

ساختمانی تاسیسات افزایش می‌یابد و برعکس، هزینه هائی که ممکن است خراب شدن تأسیسات را ببار آورد کاهش می‌یابد. در شکل ۱۷-۲۱ نمونه‌ای از روند تغییرات این دو نوع هزینه‌نشان داده شده است. مجموع این هزینه‌ها نیز که بصورت منحنی رسم شده است با بزرگ شدن دوره برگشت در ابتدا کاهش و سپس افزایش پیدا می‌کند نقطه حداقل که در واقع منطبق بر محل تلاقی دو منحنی هزینه می‌باشد نقطه‌ای است که بهترین دوره برگشت را برای طراحی در اختیار ما قرار می‌دهد.

جدول ۱۷-۲۱ دورهٔ برگشت جهت طراحی نسبت به عمر پروژه و ریسک مجاز

ریسک مجاز	عمر اقتصادی طرح (سال n)							
	1	2	5	10	20	25	50	100
0.99	1.01	1.11	1.66	2.71	4.86	5.95	11.4	22.2
0.95	1.05	1.29	2.22	3.86	7.16	8.85	17.2	33.9
0.90	1.11	1.46	2.71	14.86	9.19	11.04	22.02	43.09
0.75	1.33	2.00	4.13	7.73	14.9	18.6	36.6	72.6
0.50	2.00	3.41	7.73	14.9	29.4	36.6	72.6	145
0.33	3.00	5.45	12.9	25.2	49.9	62.1	124	247
0.25	4.00	7.46	17.9	35.3	70.0	87.3	174	346
0.20	5.00	9.47	22.9	45.3	90.1	113	225	449
0.10	10.0	19.05	48.0	95.4	190	238	475	950
0.05	20.0	39.5	98	195	390	488	975	1950
0.02	50.0	99.0	248	495	990	1238	2476	4951
0.01	100	199.5	498	995	1990	2488	4977	9953



شکل ۱۷-۲۱ متوسط هزینه‌های سالانه برای دوره‌های برگشت مختلف

مسائل

۱۷-۱ یک پل بر اساس سیل با دوره برگشت ۵۰ ساله طراحی شده است. اگر عمر این پل ۱۰۰ سال در نظر گرفته شود احتمال اینکه در طی این ۱۰۰ سال یکبار چنین سیلی رخ دهد چقدر است.

(جواب - ۲۷ درصد)

۱۷-۲ سرریز یک سد برای سیل‌های با دوره برگشت ۵۰ سال ساخته شده است. چنانچه عمر سد ۱۰۰ سال در نظر گرفته شود با چه احتمالی در طول عمر سد سیل طرح حداقل یکبار یا بیشتر رخ خواهد داد.

(جواب - ۸۷ درصد)

۱۷-۳ برای سرریز یک سد عمری معادل ۷۵ سال در نظر گرفته شده است. اگر ریسک خراب شدن سرریز در اثر سیلابها ۵ درصد در نظر گرفته شود این سر ریز با چه دوره برگشتی باید طراحی شود.

(جواب - ۱۴۶۰ سال)

۱۷-۴ بر اساس جداول ضریب فراوانی و احتمال در توزیع‌های مختلف کاغذهای احتمال مربوط را بسازید.

منابع برای مطالعه بیشتر

1- Chow, V.T., *Handbook of applied hydrology,* Mac Graw Hill, New York, 1964.

2- Duncan, A., *Quality control and statistics.* Homewood, III, Richard D. Irwin, 1959.

3- Hazen, A., *Flood flow,* John Wiley, New York, 1930.

4- kendall, G. R., *Statistical analysis of extreme values,* NRCC, Canada, Nov. 4, 1959.

5- Kite, G. W., *Frequency analysis in hydrology,* water Res. Pub. Colorado. USA, 1985.

6- Sharp J. and P. Sawden, *BASIC hydrology,* Butterworth, London, 1984.

7- Shaw, E., *Hydrology in practice,* Van Nostrand Reinhold, London, 1988.

8- Viessman, W. et al, *Introduction to hydrology,* IEP, New York, 1972.