دانشگاه شهید چمران اهواز

دانشکده کشاورزی

گروه علوم و مهندسی خاک

هیدرولوژی

آمار و احتمالات در هیدرولوژی

تهیه و تنظیم

حیدر غفاری

۲-۱- مقدمه

هر گونه طرح و برنامه‌ریزی در حوزه‌های آبریز باید براساس تجزیه و تحلیل داده‌ها و اطلاعات مربوط به هیدرولوژی آن حوزه باشد. داده‌های هیدرولوژیکی که اصطلاحاً به آنها متغیرهای هیدرولوژیکی هم گفته می‌شود به مجموعه اطلاعاتی اطلاق می‌شود که به نحوی بر سیکل هیدرولوژی موثرند. متغیرهای هیدرولوژیکی را می‌توان در چهار گروه دسته بندی کرد :

الف- متغیرهای اقلیمی

ب- متغیرهای هیدرولیکی

ج- متغیرهای مربوط به خاک، پوشش گیاهی و زمین شناسی

د- متغیرهای مربوط به فیزیوگرافی حوزه

۲-۲- سری‌های هیدرولوژیکی

برای هرگونه تجزیه و تحلیل آماری وجود مجموعه‌ای از متغیرهای هیدرولوژیکی که اندازه‌گیری و ثبت شده باشند ضروری است. به این مجموعه آماری، سری‌های هیدرولوژیکی گفته می‌شود. سری‌های هیدرولوژیکی سه گونه‌اند:

الف- سری‌های کامل[1]

ب- سری‌های جزئی[1]

ج- سری‌های حد نهایی حداکثر و حداقل[2]

- در سری‌های کامل، کلیه داده‌ها مستقل از مقدار و زمان وقوع در نظر گرفته می‌شوند (حجم داده‌ها بسیار زیاد است و محاسبات را مشکل می‌سازد).

- در سری‌های جزئی، فقط ارقامی که بالاتر یا پایین‌تر از یک حد مشخص باشند، درنظر گرفته می‌شوند.

- در سری‌های حد نهایی، بالاترین و یا پایین‌ترین مقدار یک پارامتر که در یک دوره زمانی مشخص رخ داده است، مورد تجزیه و تحلیل قرار می‌گیرد. محدودیت این سری‌ها در این است که برای هر سال فقط یک رقم منظور می‌شود، حال آن که ممکن است دومین سیل بزرگی که در یک سال اتفاق می‌افتد از شدیدترین سیلی که در سال دیگر اتفاق می‌افتد بیشتر باشد.

متغیرهایی که در سری‌های هیدرولوژیکی مورد تجزیه و تحلیل قرار می‌گیرند ممکن است به لحاظ ماهیت از دو نوع زیر باشند:

۱- متغیرهای پیوسته[۳]: پارامترهایی هستند که مقدارشان می‌تواند هر عدد صحیح یا اعشاری باشد. مانند مقدار بارندگی، درجه حرارت، فشار هوا

۲- متغیرهای ناپیوسته[۱] (که متغیرهای جدا یا گسسته نیز نامیده می‌شوند): پارامترهایی هستند که مقدارشان فقط اعداد صحیح بوده و در واقع کمیت آنها با فراوانی وقوعشان سنجیده می‌شود مانند تعداد روزهای بارانی سال، تعداد روزهای یخبندان در سال، پوشش ابر و جهت باد

۲-۳- تحلیل فراوانی‌ها

فراوانی وقوع عبارتست از تعداد دفعاتی که یک پارامتر مشخص در مدت زمان معین اتفاق می‌افتد. چنانچه داده‌های مربوط به سال‌های آماری را بدون توجه به سال وقوع به ترتیب صعودی یا نزولی ردیف کنیم احتمال وقوع هر یک از داده‌ها را می‌توانیم از یک رابطه تجربی مانند $P=\frac{m}{n+1}$ به دست آوریم که در این رابطه P= احتمال وقوع؛ m= شماره ردیف و n= تعداد کل داده‌ها می‌باشد. همچنین در جدول ۲-۱ تعدادی از روابط احتمال تجربی که جهت محاسبه احتمال وقوع، وقایع هیدرولوژیکی پیشنهاد شده-اند ارائه شده است.

دوره بازگشت که عکس احتمال وقوع است عبارت است از تعداد سال‌هایی که به طور متوسط بین وقوع دو حادثه مشابه وجود دارد.

$$T = \frac{1}{P}$$

جدول ۲-۱- روابط پیشنهادی جهت محاسبه احتمال وقوع وقایع هیدرولوژیکی

ردیف	نام رابطه	سال	احتمال	کشور
۱	کالیفرنیا[۲]	۱۹۲۳	$\frac{m}{n}$	آمریکا
۲	هیزن[۳]	۱۹۳۰	$\frac{2m-1}{2n}$	فرانسه

۲-۴- خواص احتمال

همان‌گونه که گفته شد وقایع هیدرولوژیکی می‌توانند به صورت گسسته در نظر گرفته شوند اما اغلب متغیرهایی پیوسته می‌باشند، با وجود این خصوصیات توزیع‌های احتمالاتی برای موارد گسسته $P_r(x)$ و پیوسته $f(x)$ مشابه هستند:

$$P_r(x) \quad \text{یا} \quad f(x) \geq 0$$

$$\sum_{all\,x} P_r(x) = 1 \quad \text{یا} \quad \int_{-\infty}^{+\infty} f(x)dx = 1$$

۲-۷- توابع احتمالی

به طور کلی داده‌های هیدرولوژیکی می‌توانند جزء یکی از گروه‌های زیر باشند:

۱- توابع احتمالاتی که مبتنی بر گشتاورها هستند.

۲- توابع احتمالاتی که مبتنی بر گشتاورها نیستند.

۳- توابع پلکانی

۴- سری شدت‌های بیش از حد معین

اغلب کارهای انجام شده در هیدرولوژی در گروه اول قرار دارد.

در ادامه ۱۵ تابع توزیع آماری معرفی شده‌است:

الف- توابع گسسته

۱- توزیع برنولی

۲- توزیع دوجمله‌ای[۱]

۳- توزیع پواسون[۲]

۴- توزیع هندسی

۵- توزیع فوق هندسی

ب- توابع پیوسته

۶- توزیع نرمال[۱]

۷- توزیع لوگ نرمال[۲]

۸- توزیع نمایی[۳]

۹- توزیع لوگ نرمال سه پارامتری[۴]

۱۰- توزیع گامبل[۵]

۱۱- توزیع پیرسون تیپ سه[۶]

۱۲- توزیع لوگ پیرسون تیپ سه[۷]

۱۳- توزیع گاما[۸]

۱۴- تبدیل باکس کاکس

۱۵- روش جدید ترسیمی

۱۷-۷ توابع توزیع احتمال

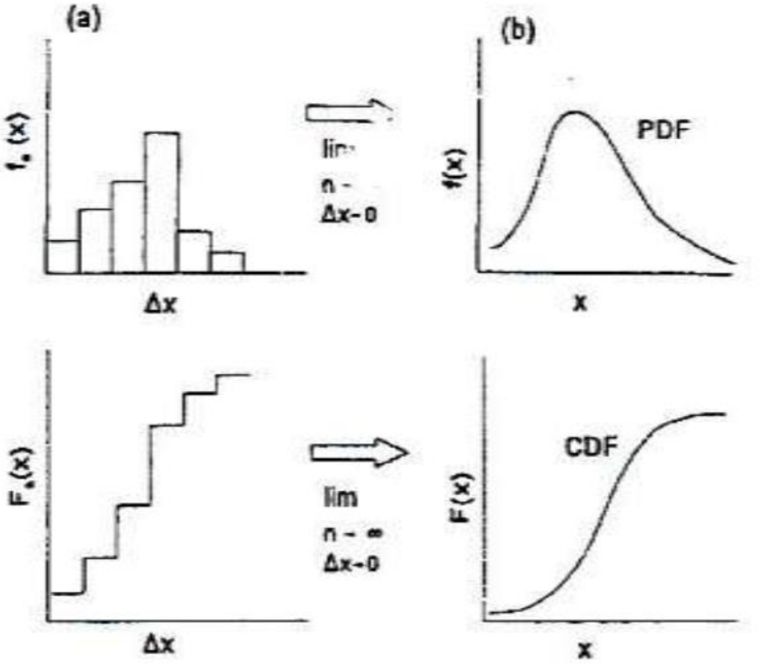
اگر چنانچه تعداد داده‌های هیدرولوژی n بوده و آن‌ها را جهت رسم هیستوگرام دسته بندی کرده باشیم و فرض کنیم که در هر دسته تعداد n_i داده قرار گیرد لذا فروانی نسبی در هر دسته n_i/n خواهد بود. بنابراین ***تابع فراوانی نسبی*** (relative frequency function) عبارت خواهد بود از:

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

از جمع کردن مقادیر فراوانی‌های نسبی با همدیگر ***تابع فراوانی تجمعی*** بدست می‌آید (cu 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 ... ی 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.

(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:

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

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

(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.

(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 ∞- تا ∞+ 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 ∞- تا ∞+ انتگرال‌گیری کنیم توزیع نرمال استاندارد (standard normal distribution) بدست می‌آید. از آن جائی که متغیر تصادفی از ∞- تا ∞+ 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:

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

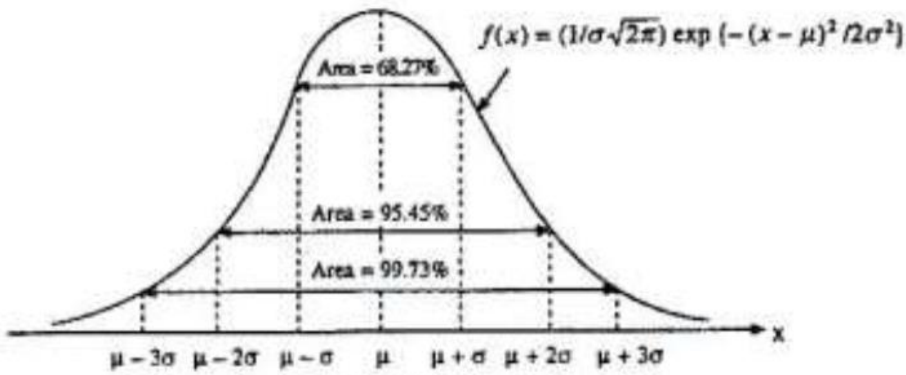
$$x = \mu + \sigma t \qquad (۴۶-۱۷)$$

معادله فوق برای داده‌های از ∞- تا ∞+ 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.



سطح زیر منحنی

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 ۱۷-۱۰ 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

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.

۱۷-۱۰ 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:

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

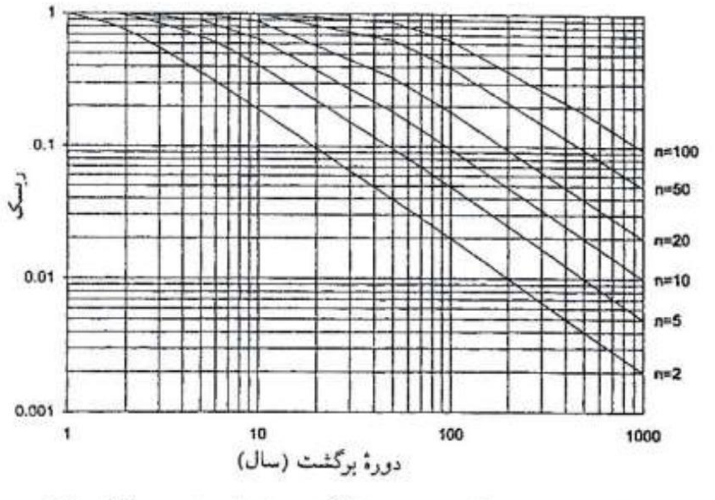
بهمین ترتیب احتمال این که واقعه مذکور در n سال آینده رخ ندهد $(1 - P)^n$ می‌باشد. بنابراین احتمال رخ دادن همین حادثه در طی n سال آینده برابر خواهد بود با:

$$J = 1 - (1 - P)^n \qquad (۱۷-۷۴)$$

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)

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 بامِلات	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.