Stochastic methods in water resources

Lecture 4: Moments, characteritics function, known probability distributions, bivariate probability distributions

Luis Alejandro Morales, Ph.D.

Universidad Nacional de Colombia

August 7, 2025

Expected value

The measures or descriptors of a random variable determine important features of their behaviour and thus describe the pmfs and pdfs. The descriptors are defined upon the pmfs o pdfs and are known as the statistics of the random variable.

Expected value

If X is a discrete random variable and $g(x)$ is a function of the random variable, the **expected value** of $g(x)$ is:

$$E\left[g(x)\right] = \sum g(x_i) p_X(x_i)$$

If X is continuos, **expected value** is:

$$E\left[g(x)\right] = \int_{-\infty}^{\infty} g(x) f_X(x)$$

These equations indicate that the expected value or average of a random variable X is computed by weighting the random variable function $g(X)$ with the pmf or the pdf. Some of the properties of the expected value operator, E, are:

$$\begin{aligned}
E(a) &= a, \text{ when } a \text{ is constant} \\
E\left[ag(x)\right] &= aE\left[g(x)\right], \text{ when } a \text{ is constant} \\
E\left[ag_1(x) + bg_2(x)\right] &= aE\left[g_1(x)\right] + bE\left[g_2(x)\right], \text{ if } a \text{ y } b \text{ are constants} \\
E\left[g_1(x)\right] &\leq E\left[g_2(x)\right], \text{ if } g_1(x) \geq g_2(x)
\end{aligned}$$

Moments

Moments

The **moments** are a family of averages or expected values of a random variable used to describe its behaviour. It is common to characterize a pmf or a pdf using its main moments. If μ_r^* represents the moment of order r around the point a, this is defined for a discrete variable as:

$$\mu_r^* = E\left[(X-a)^r\right] = \sum_{\text{todo } x_i} (x_i - a)^r p_X(x_i)$$

for a continuous variable:

$$\mu_r^* = E\left[(X-a)^r\right] = \int_{-\infty}^{\infty} (x-a)^r f_X(x) dx$$

where r is an positive number. The most known moments to describe *pmfs* or *pdfs* are the **mean** (μ_X), the **variance** ($s_X = \sigma_X^2$) and the asymmetry γ.

Mean, μ_X

The **mean** is the first moment ($r = 1$) with respect to the origin ($a = 0$). It is measure that describe the central tendecy of the random variable. The **mean** has the same units that X. This is defined as:

Discrete random variable X

$$\mu_X = E[X] = \sum_{\text{todo } x_i} x_i p_X(x_i)$$

Continuous random variable X

$$\mu_X = E[X] = \int^{\infty} x f_X(x) dx$$

Moments

Variance, σ_X^2

The **variance** is the second moment ($r = 2$) with respect to the $a = \mu_X$. The **variance** describes the variability of a random variable X around their expected value μ_X.

Discrete random variable X

$$\sigma_X^2 = E\left[(X - \mu_X)^2\right] = \sum_{i=1}^{n} (x_i - \mu_X)^2 p_X(x_i)$$

Continuous random variable X

$$\sigma_X^2 = E\left[(X - \mu_X)^2\right] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$$

The larger σ_X^2, the larger the dispersion of X aroung μ_X. According to the properties of E discussed above, one can demostrate that:

$$\sigma_X^2 = E\left[(X - \mu_X)^2\right] = E[X^2] - (E[X])^2$$

The **standard deviation**, σ_X, is defined as:

$$\sigma_X = \sqrt{\sigma_X^2}$$

where σ_X has the same units of X and thus facilitate the understanding of the dispersion of X around μ_X. The relative measure of dispersion with respect to μ_X is the **coefficient of variation** defined as:

$$V_X = \frac{\sigma_X}{\mu_X}$$

Moments

Asymmetry, γ

The **asymmetry**, γ, is the third moment ($r = 3$) around the mean ($a = \mu_X$). The definition is:

Discrete random variable X

$$\gamma = E\left[(X-\mu_X)^3\right] = \sum_{i=1}^{n}(x_i-\mu_X)^3 p_X(x_i)$$

Continuous random variable X

$$\gamma = E\left[(X-\mu_X)^3\right] = \int_{-\infty}^{\infty}(x-\mu_X)^3 f_X(x)dx$$

γ is a measure of the pmf or pdf symmetry. Based on these equations, the nondimensional asymmetry coefficient is defined as:

$$\gamma_1 = \frac{E\left[(X-\mu_X)^3\right]}{\sqrt{\left[E\left[(X-\mu_X)^2\right]\right]^3}} = \frac{E\left[X^3\right] - 3E\left[X^2\right]\mu_X + 2\mu_X^3}{\left(E\left[X^2\right]-\mu_X^2\right)^{\frac{3}{2}}}$$

As you can see in the equation above, the numerator is the central moment of third order ($r = 3$) and the denominator is σ_X raised to the power of 3. Note that γ_1 have the same sign that γ. For symetrical distribution, $\gamma \approx 0$.

Moments

Quantiles, q

The q-esimo quantile, q, of a random variable X, is defined as the smallest number ξ that satisfy the inequality $F_X(\xi) \geq q$. For a continuos variable, ξ sastisfy $F_X(\xi) = q$ and the quantile is the value of X that is exceeded with a probability of $(1 - q)$. Convesely, the quantile can be written as $q = x(F)$ or $q = x[F_X(x)]$. Accordingly, the mediana is the 0.5 quantile, written as $\xi_{0.5}$. A quantile q can be defined as $\xi_q = F_X^{-1}(q)$. The most common q are the quartiles $\xi_{0.25}$, $\xi_{0.5}$ and $\xi_{0.75}$.

Daily evaporation

Suppose that the pdf of the evaporation V for any day of the year is defined as:

$$f_V(v) = \begin{cases} 0.125, \text{ if } 0.5 \leq v \leq 8.5 \; mm/day \\ 0 = \text{ otherwise} \end{cases}$$

Estimate the main moments.

Mean μ_V

$$\mu_V = E[V] = \int_{-\infty}^{\infty} v f_V(v) dv$$
$$= \int_{0.5}^{8.5} 0.125 v dv$$
$$= 0.0625 v^2 \Big|_{0.5}^{8.5} = 4.5 mm/day$$

Variance σ_V^2

$$\sigma_V = E\left[(v - \mu_V)^2\right]$$
$$= \int_{-\infty}^{\infty} (v - \mu_V)^2 f_V(v) dv$$
$$= 0.125 \int_{0.5}^{8.5} (v - 4.5)^2 dv$$
$$= \frac{0.125}{3} \left[(v - 4.5)^3\right] \Big|_{0.5}^{8.5} = 5.33 mm^2/day^2$$

Moments

Daily evaporation

Standard deviation σ_V

$$\sigma_V = \sqrt{\sigma_V^2} = \sqrt{5.33} = 2.31 mm/day$$

Quantile ξ

From the $F_V(v) = 0.125v - 0.0625$ and regarding that $\xi_q = F_V^{-1}(q)$, one has that $\xi_q = 8(q + 0.0625)$. So $\xi_{0.1} = 1.3$ and $\xi_{0.5} = 4.5$ (the mediana). The interquartile range is $q_r = \xi_{0.75} - \xi_{0.25} = 6.5 - 2.5 = 4.$

Assymetry γ

$$\gamma = E\left[(v - \mu_V)^3\right] = \int_{-\infty}^{\infty} (v - \mu_V)^3 f_V(v) dv$$

$$= 0.125 \int_{0.5}^{8.5} (v - 4.5)^3 dv$$

$$= \frac{0.125}{4} \left[(v - 4.5)^4\right] \Big|_{0.5}^{8.5} = 0$$

This indicate that $f_V(v)$ is simetric with respecto to μ_V.

Generating functions

A **generating function** is a convenient way to represent a sequence such as the sequence of moments. The function, usually of a quantity t, is expanded as power series to give the values of the moments as the coefficients.

Moment-generating function

Moments define a pmf or a pdf. For some distributions, it is possible to define a **moment-generating function**, **mgf** that exist for a domain $-\varepsilon < t < \varepsilon$. The mgf is defined as:

$$M_X(t) = E[e^{tX}]$$

Mathematically, e^{tX} can be expanded in powers of t about of zero as a Maclauring's series. Replacing in the equation above and getting the expected values for the expansion coefficients:

$$M_X(t) = E[e^{tX}] = E\left[1 + Xt + \frac{1}{2!}(Xt)^2 + \cdots\right] = 1 + \mu_1 t + \frac{1}{2!}\mu_2 t^2 + \cdots$$

one can get the moments of the distribution. Alternatively for:

discrete x	**continuous** x
$M_X(t) = \sum e^{tx_j} p_x(x_j)$, for all possible x_j	$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_x(x) dx$

The moment of order m can be obtained by deriving $M_X(t)$ with respect to t and evaluating the derivative for $t = 0$, so:

$$\frac{d^m M_X(t=0)}{dt^m} = \int_{-\infty}^{\infty} x^m f_X(x) dx = E[X^m]$$

Generating functions

Moment-generating function

For instance, the first moment $(m = 1)$ is obtained as:

$$\frac{dM_X(t=0)}{dt} = \left[\int_{-\infty}^{\infty} xe^{tx} f_X(x)dx\right]_{t=0} = \int_{-\infty}^{\infty} xf_X(x)dx = E[X]$$

The second moment, $(m = 2)$, with respect to zero is:

$$\frac{d^2M_X(t=0)}{dt^2} = \int_{-\infty}^{\infty} x^2 f_X(x)dx = E[X^2]$$

Time between rainfall events

The time (e.g days, hours) between rainfall events is a continuous variable X exponentially distributed as $f_X(x) = \lambda e^{-\lambda x}$ for $0 < X < \infty$. The moment-generating function is:

$$M_X(t) = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}dx = \frac{\lambda}{t-\lambda} e^{x(t-\lambda)}\Big|_0^{\infty} = \frac{\lambda}{\lambda - t}$$

As $t < \lambda$, the function above evaluated in ∞ is zero. The equation above gives the moment-generating function. The firs moment $(m = 1)$ is estimated as:

$$\frac{dM_X(t=0)}{dt} = E[X] = \mu_X = \frac{\lambda}{(\lambda - t)^2}\Big|_{t=0} = \frac{1}{\lambda}$$

Time between rainfall events

The second moment is:

$$\frac{d^2 M_X(t=0)}{dt^2} = E[X^2] = \frac{2\lambda}{(\lambda - t)^3}\Big|_{t=0} = \frac{2}{\lambda^2}$$

Following the definition of the variance σ_X^2:

$$\sigma_X^2 = E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Factorial moment-generating function

For some discrete random variable, it is convenient to apply a **factorial moment-generating function** defined as $E[t^X]$. For this function, the condition of interest is $t = 1$ instead of $t = 0$ (used in the moment-generating function). If $E[t^X]$ exist, the mth derivative for $t = 1$ is the mth-order factorial moment of X.

Number of rainy days

Number of rainy days

The number of rainy days in a year is a discrete random variable X and it is described by a Poisson distribution $p_X(x) = \frac{\nu^x e^{-\nu}}{x!}$, where $x = 0, 1, 2, \cdots, \nu > 0$. As X has a factorial mgf:

$$E[t^X] = \sum_{x=0}^{\infty} \frac{t^x \nu^x e^{-\nu}}{x!} = e^{-\nu} \sum_{x=0}^{\infty} \frac{(t\nu)^x}{x!} = e^{-\nu} e^{t\nu} = e^{\nu(t-1)}$$

The first and second derivative of $E[t^X]$ with respect to t are:

first derivative

$$\frac{dE[t^X]}{dt} = \nu e^{\nu(t-1)}\Big|_{t=1} = \nu$$

second derivative

$$\frac{d^2E[t^X]}{dt^2} = \nu^2 e^{\nu(t-1)}\Big|_{t=1} = \nu^2$$

Similarly, The first and second derivative of $E[t^X]$ with respect to t are using the E operator:

first derivative

$$\frac{dE[t^X]}{dt} = E\left[Xt^{(X-1)}\right]_{t=1} = E[X]$$

second derivative

$$\frac{d^2E[t^X]}{dt^2} = E\left[X(X-1)t^{(X-2)}\right]_{t=1} = E[X^2] - E[X]$$

Equaling the above equations:

$$E[X] = \nu \qquad E[X^2] - E[X] = \nu^2$$

Accordingly:

$$\sigma_X^2 = E[X^2] - (E[X])^2 = \nu^2 + \nu - \nu^2 = \nu$$

Characteristic functions

Characteristic functions

A **characteristic function** of X is an alternative function useful when the mgf does not provide estimates of moments. It is defined as:

$$\phi_X(t) = E[e^{itx}] = M_X(it)$$

discrete x

$$M_X(it) = \sum e^{itx_j} p_x(x_j), \text{ for all possible } x_j$$

continuous x

$$M_X(it) = \int_{-\infty}^{\infty} e^{itx} f_x(x) dx$$

donde $i = \sqrt{-1}$. For instance for a continuous X, the characteristic function is similar to the Fourier transform, so the inverse is:

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{itx} \phi_X(t) dt$$

This equation means that if two random variable X and Y have the same characteristic function, they are identically distributed.
The characteristic function can be used to calculate the moments as:

$$\mu^r = E[X^r] = \frac{1}{i^k} \frac{d^k \phi_X(t)}{dt^k}\Big|_{t=0}$$

Expanding the caracteristic function $\phi_X(t)$ in Taylor's series around $t = 0$, we have:

$$e^{itX} = 1 + itX + \frac{1}{2}(itX)^2 + \frac{1}{6}(itX)^3 + \cdots$$

Characteristic functions

Characteristic functions

Taking expectation in both sides:

$$\begin{aligned}\phi_X(t) = E[e^{itX}] &= 1 + itE[X] + \frac{1}{2}(it)^2 E[X^2] + \frac{1}{6}(it)^3 E[X^3] + \cdots \\ &= 1 + it\mu + \frac{1}{2}(it)^2\mu^2 + \frac{1}{6}(it)^3\mu^3 + \cdots\end{aligned}$$

In a compact manner:

$$\phi_X(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!}\mu_r$$

This equation indicate that if the moments exist and it converges, the pdf of X is completely defined. This is de case of most pdfs in practice. If there are two independent random variables X and Y, the characteristic function is:

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$$

It follows that if we have the sum of M indentically distributed random variables $Z = \sum_{i=1}^{M} X_i$, the characteristic function of Z is:

$$\phi_Z(t) = [\phi_X(t)]^M$$

For a variable $Y = a + bX$, the caracteristic function is:

$$\phi_Y(t) = e^{itb}\phi_X(at)$$

Characteristic functions

Characteristic functions

Taking logarithms for equation $\phi_Z(t) = [\phi_X(t)]^M$:

$$K_Z(t) = MK_X(t)$$

where the functions $K_Z(t) = \ln \phi_Z(t)$ and $K_X(t) = \ln \phi_X(t)$ are the **cumulant functions** of Z and X, respectively. The series expansion for the cumulant function is:

$$K_Z(t) = \sum_{n=1}^{\infty} \frac{(it)^n}{n} \kappa_n$$

where the **cummulants**, $\kappa_n = \frac{1}{i^n} \frac{d^n K_Z(t)}{ds^n}$. The cumulants are related to the pdf moments and viseversa so:

$$\begin{aligned}
\kappa_1 &= \mu_1 \\
\kappa_2 &= \mu_2 - \mu_1^2 = \sigma^2 \\
\kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 3\mu_1^3 \\
\kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4
\end{aligned}$$

Estimation of parameters

Statistical inference serves to get an estimate of the experiment parameters based on a random sample from the population. In statistical inference is assumed that the distribution of the population is known. The **estimator** is thus a method to obtain a parameter value based on the sample and can be bias. This methods of parameter estimation are known as **point estimation**. In summary, the problem of statistical inference is to find the parameters of a probability distribution (or parametric distribution) that fit the best the sample data. The most known methods (estimators) to perform that are treated here :the method of moments, Method of probability weighted and L-moments and Maximum Likelihood.

The method of moments

The **method of moments** is the most know analytical estimator. Given a pdf of the population and applying the concept of moments, one can get analytical expressions for the moments (.e.g μ_X and σ_X^2) as function of the pdf parameters. For instance, in the case of a pdf with two parameters (α_1 y α_2), the concept of moments is used to obtain two equations, one for μ_X and the other for σ_X^2. An estimate of $\hat{\mu}_X$ and $\hat{\sigma}_X^2$ is calculated based on the sample data. This two estimates are replaced in the two equations to get the pdf parameters. One of the problem with the method is that the parameter estimates may be biased, that is, if for m different samples of size n, different pdf parameters are estimated, the average of the parameter values does not converge to the real value of the parameter. Also, the estimator can be ineficient and an efficient estimator has the smallest variance among all posible estimator.

Estimation of parameters

Method of probability weighted and L-moments

For a random variable X with cdf $F_X(x)$, the **probabilistic weighted moments (pwms)** are defined as:

$$M_{ijk} = E\left[X^i\{F_X(x)\}^j\{1 - F_X(x)\}^k\right] = \int_0^1 x(F)^i F^j (1-F)^k dF$$

where $x(F)$ is the quantile or inverse cdf function of X, and j and k take values as $0, 1, \cdots, m$, where m is the number of samples. For $j = k = 0$ this equation provides the moment of order i about zero. In the application of pwms, it is often convenient to $i = 1$, and $j = 0$ or $k = 0$.

Extreme storms

Extreme values such as flood discharges, precipitation over threshold, wind waves, etc, usually follow a extreme value distribution. If X is a extreme random variable, the cdf is given by:

$$F_X(x) = \exp\left[-\exp\left(-\frac{x-b}{a}\right)\right]$$

where a and b are the cdf parameters. The quantile function $X(F)$ is:

$$\ln y = -\exp\left(-\frac{x-b}{a}\right)$$

$$\ln(-\ln y) = -\frac{x-b}{a}$$

$$x = b - a\ln(-\ln y)$$

where $y \equiv F_X(x)$.

Estimation of parameters

Extreme storms

Applying the expression for the pwms when $i = 1$ and $k = 0$, one has:

$$M_{ijk} = M_j = \int_0^1 (b - a\ln(-\ln y))\, y^j\, dy = b\int_0^1 y^j\, dy - a\int_0^1 y^j \ln(-\ln y) dy$$

The solution of the first integral is easy $b\int_0^1 y^j\, dy = \frac{b}{1+j}$. The solution of the second integral is complicated, this is $a\int_0^1 y^j \ln(-\ln y) dy = \frac{-a[\ln(1+j)+n_e]}{1+j}$. $n_e \approx 0.57721$ is the Euler's number. The solution of the integral is:

$$M_j = \frac{b}{1+j} + \frac{a\,[\ln(1+j) + n_e]}{1+j}$$

For pwms for $j = 0$ and $j = 1$ are:

$j = 0$

$$M_0 = b + an_e$$

$j = 1$

$$M_1 = \frac{b + a(\ln 2 + n_e)}{2}$$

Expressing the equations above in terms of a and b:

$$a = \frac{2M_1 - M_0}{\ln 2}$$

$$b = M_0 - an_e$$

The problem is how to estimate M_0 and M_1 to calculate a and b.

Estimation of parameters

Extreme storms

Suppose that there is a sample of X that take values $(x_0, x_1, \cdots, x_n)$ where n is sample size. Experimentally, pwms are calculated as

$$M_j = \frac{1}{n} \sum_i p_i^j x_i$$

where p_i is the probability of x_i. Accordingly:

$j = 0$

$$M_0 = \hat{\mu}_X = \frac{1}{n} \sum_i x_i \text{ (mean)}$$

$j = 1$

$$M_1 = \frac{1}{n} \sum_i p_i x_i \text{ (weighted mean)}$$

Suppose that we have the following table with the fourteen maximum precipitation depth (x_i (mm)) for a 3-h storm duration for a period of time. The table shows the ordered precipitation and the probabilities of each value p_i.

Order, i	$x_{(i)}$ (mm)	p_i	$P_i\ x_{(i)}$ (mm)
1	32.3	0.046	1.500
2	41.2	0.118	4.856
3	49.9	0.189	9.445
4	51.6	0.261	13.453
5	56.0	0.332	18.600
6	56.7	0.404	22.883
7	63.8	0.475	30.305
8	66.5	0.546	36.338
9	80.6	0.618	49.799
10	90.1	0.689	62.105
11	114.2	0.761	86.874
12	122.3	0.832	101.771
13	157.3	0.904	142.132
14	188.7	0.975	183.983

[a] The corresponding plotting positions, p_i, and the product $p_i x_{(i)}$ are given in order to use the probability weighted moments method for evaluating the EV1 distribution.

Note that the values of p_i correspond to plotting positions for the ordered values x_i. According to the values in the table, $M_0 = 83.7$ mm and $M_1 = 54.6$ mm. The estimated values of $\hat{a}$ and $\hat{b}$ are:

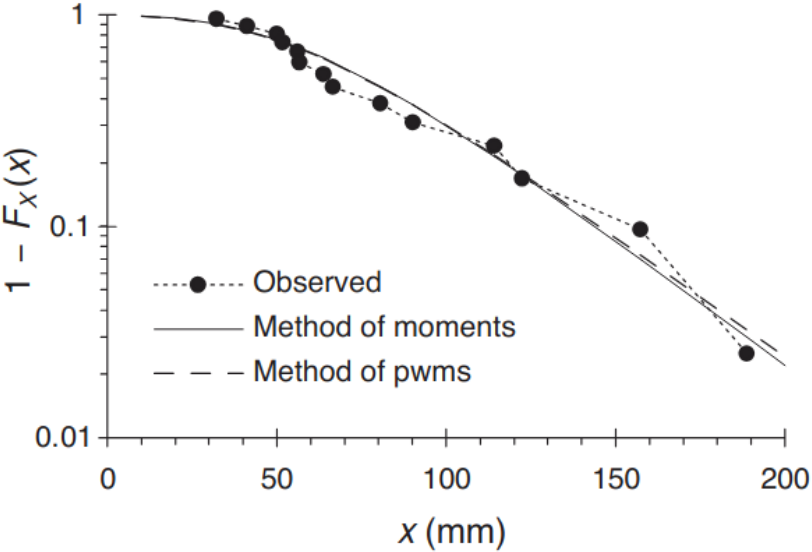
$$\hat{a} = \frac{2M_1 - M_0}{\ln 2} = \frac{2x54.6 - 83.7}{\ln 2} = 36.8 \text{ mm}$$

$$\hat{b} = M_0 - an_e = 83.7 - 36.8x0.5772 = 62.4 \text{ mm}$$

Estimation of parameters

Extreme storms

The following figure shows the probability of exceedence ($Pr[X > x] = 1 - F_X(x)$) for the observed data, the theoretical extreme value distribution whose coefficients (a y b) were estimated using the methods of moments and the pwms.



Estimation of parameters

Maximum likelihood method

For a given random variable X, a sample of size n is taken from the space space, and the pdf is $f_X(x)$. The **likelihood function** of θ, where θ is vector of m parameters, is represented as:

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i|\theta)$$

The objective is to find the vector θ that maximize $L(\theta)$ for a given X sample. It is thus necessary to obtain the m partial derivative of $L(\theta)$ with respect to each θ parameter. Derivatives are equated to zero and solved to find the **maximum likelihood (ML)** estimators ($\hat{\theta}$) of θ.

Flood exeedance

Suppose that X is a discrete random variable that determine the exceedance of a discharge threshold. Accordingly, X is equal to 1 where $X > x_{threshold}$ (flood occurrence) and equal to zero where $X \leq x_{threshold}$ (no occurrence of flood, complementary event). After the analysis of gauge streamflow data, the probability of $X = 1$ is equal to p and the probability of $X = 0$ is $(1 - p)$. If the pmf of X is a Bernoulli distribution:

$$P_X(x_j) = Pr[X = x_j] = p^{x_j}(1-p)^{1-x_j}, \text{ for } x_j = 0, 1$$

Note that p is the parameter of the *pmf*. The likelihood function for n trials or outcomes is:

$$L(p) = \prod^{n} p^{x_j}(1-p)^{1-x_j}, \text{ for } x_j = 0, 1$$

Estimation of parameters

Flood exeedance

A fundamental identity indicate that $\ln\left(\prod_{i=1}^{n} a_i\right) = \sum_{i=1}^{n} \ln a_i$. Following this:

$$\ln L(p) = \ln\left(\prod_{j=1}^{n} p^{x_j}(1-p)^{1-x_j}\right) = \sum_{j=1}^{n} \ln\left(p^{x_j}(1-p)^{1-x_j}\right)$$

Then:

$$\ln L(p) = \ln p \sum_{j=1}^{n} x_j + \ln(1-p)\sum_{j=1}^{n}(1-x_j)$$

The derivative is:

$$\frac{d\left(L(p)\right)}{dp} = \frac{1}{p}\sum_{j=1}^{n} x_j - \frac{1}{1-p}\sum_{j=1}^{n}(1-x_j) = \frac{\sum_{j=1}^{n} x_j}{p} - \frac{n-\sum_{j=1}^{n} x_j}{1-p}$$

Equaling to zero and solving for p:

$$\hat{p} = \frac{\sum_{j=1}^{n} x_j}{n}$$

where $\hat{p}$ is and estimator of p.

The maximum likelihood method is the most implemented method. However, large samples are required to prevent the estimator becomes unbiased. This estimator does not have a low variance in comparison with others. Sometimes the estimators can be obtained analytically so numerical methods are needed.

Multiple random variables

- So far, only single random variables have been considered belonging to a population.
- To describe individual random variables, univariate distributions have been considered.
- In these section, experiments where two or more random variables occur simultaneously are studied.
- Variables are studied jointly and their distribution are of the multivariate type.
- For multiple variables, probabilities laws are described by joint probability mass or density functions.
- An example of a continuous bivariate distribution is given by the mean hourly wind speed and wind direction recorded by a weather station. Note that both random variables occurred simultaneously.

Joint probability distribution of discrete variables

Joint probability mass function

Given two discrete random variables X y Y, the joint (bivariate) **pmf** is given by the intersection probability $p_{X,Y}(x,y) = Pr[(X=x)\cap(Y=y)]$. According to the proability axioms, $\sum_{\text{all } x_i}\sum_{\text{all } y_j} p_{X,Y}(x_i,y_j) = 1$. The joint **cdf** is given by $F_{X,Y}(x,y) = Pr[(X\leq x)\cap(Y\leq y)] = \sum_{x_i\leq x}\sum_{y_j\leq y} p_{X,Y}(x_i,y_j)$. For n randon variables X, the joint pmf is given by:

$$p_{X_1,X_2,\cdots,X_n} = Pr[(X_1 = x_1)\cap(X_2 = x_2)\cap\cdots\cap(X_n = x_n)]$$

and the joint cdf is given by:

$$\begin{aligned}F_{X_1,X_2,\cdots,X_n} &= Pr[(X_1\leq x_1)\cap(X_2\leq x_2)\cap\cdots\cap(X_n\leq x_n)]\\ &= \sum_{x_{1_i}\leq x_1}\sum_{x_{2_j}\leq x_1}\cdots\sum_{x_{n_k}\leq x_n} p_{X_1,X_2,\cdots,X_n}(x_{1_i},x_{2_j},\cdots,x_{n_k})\end{aligned}$$

Wind records

In an urban area, two different weather stations, whose precisions are different, measure wind speed. Engineers are interested to measure the number of days per year that wind speed, in each station, surpases an speed threshold. As these winds are liable to cause infrastructure damages, for design purposes, it is needed to estimate the joint probability of the number of days per year where wind speed is $>$ 60 km/h estimated using the two instruments. The following table shows the pmf for the variables measured at the accurate station X and the less accurate station Y.

Joint probability distribution of discrete variables

Wind records

	$Y = 0$	$Y = 1$	$Y = 2$	$Y = 3$	$p_X(x)^a$
$X = 0$	0.2910	0.0600	0.0000	0.0000	0.3510
$X = 1$	0.0400	0.3580	0.0100	0.0000	0.4080
$X = 2$	0.0100	0.0250	0.1135	0.0300	0.1785
$X = 3$	0.0005	0.0015	0.0100	0.0505	0.0625
$p_Y(y)^a$	0.3415	0.4445	0.1335	0.0805	$\Sigma = 1.0000$

[a] The entries in the last column and bottom row are the respective marginal probabilities.

Note that this table also show the **marginal pmf** of X and Y. Suppose that one want to judge the accuracy of Y, the probability that $(X = Y)$ (event A) can be calculated as:

$$Pr[A] = \sum_{\text{all } x_i} p_{X,Y}(x_i, y_i) = p_{X,Y}(0,0) + p_{X,Y}(1,1) + p_{X,Y}(2,2) + p_{X,Y}(3,3) = 0.813$$

Joint probability distribution of discrete variables

Conditional probability mass function

Given two discrete random variables X and Y whose joint pmf is given by $p_{X,Y}(x,y)$, the **conditional probability mass function** given that a value of Y equal to y_j is:

$$p_{X|Y}(x|y_j) = Pr[X = x|Y = y_j] = \frac{Pr[(X = x) \cap (Y = y_j)]}{Pr[Y = y_j]}$$

$$= \frac{p_{X,Y}(x, y_j)}{\sum_{\text{all } x_i} p_{X,Y}(x_i, y_j)} = \frac{p_{X,Y}(x, y_j)}{p_Y(y_j)} \quad \text{, for all } j$$

Accorging to the probability axioms, $0 \leq p_{X,Y}(x|y_j) \leq 1, \quad \text{for all } j$ and $\sum_{\text{all } x_i} p_{X|Y}(x_i, y_j) = 1, \quad \text{for all } j$. Other conditional probability distribution are derived, for instance, when one wants to know the probability distribution of X when $Y \geq y$. So:

$$p_{X|Y \geq y} \equiv Pr[X = x|Y \geq y] = \frac{\sum_{y_j \geq y} p_{X,Y}(x, y_j)}{\sum_{y_j \geq y} p_Y(y_j)}$$

Engineers are aware that almost all observed events are conditioned by the occurrence of other events. This is why, in practice, conditional probabilities are more easily obtained than joint pmfs. Accordingly, the joint pmf can be obtained using the conditional pmf and the corresponding marginal pmf as follows:

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y) = p_{Y|X}(y|x)p_X(x)$$

Wind records

According to the table above, if $Y = 1$, the joint probabilities $p_{X,Y}(x, 1)$ are given by 0.0600, 0.3580, 0.0250 and 0.0015 for $x = 0, 1, 2$ and 3. The sum of these probabilities, $p_Y(Y = 1) = 0.4445$. The conditional pmf, $p_{X|Y}(x|1)$ is given by 0.1350, 0.8054, 0.0562 and 0.0034, for $x = 0, 1, 2$ and 3. Note that the sum of $p_{X|Y}(x|1)$ is 1.

Marginal probability mass function

For multiples random variables, if all variables are disregarded apart from the single variable X_i, the **marginal pmf** of X_i can be obtained from the joint pmf. For instance, for a bivariate joint pmf, the marginal pmf of X (or for Y) is:

$$p_X(x) \equiv Pr[X = x] = \sum_{\text{all } y_j} Pr[X = x|Y = y_j]Pr[Y = y_j] = \sum_{\text{all } y_j} p_{X,Y}(x, y_j)$$

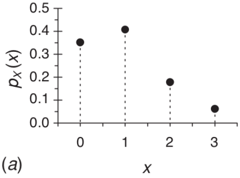
The cdf can be obtained from this equation as follow:

$$F_X(x) \equiv Pr[X \leq x] = \sum_{x_i \leq x} \sum_{\text{all } y_j} p_{X,Y}(x_i, y_j)$$

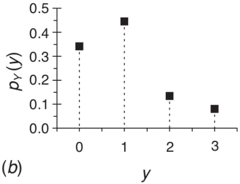
Joint probability distribution of discrete variables

Wind records

From the table of wind records, suppose that the engineer want to know the $Pr[X = 0]$, so applying marginal pmf equation $p_X(0) = Pr[X = 0] = \sum_{y=0}^{3} p_{X,Y}(0, y) = 0.2910+0.0600+0.0000+0.0000 = 0.3510$. The $Pr[X > 0]$ can be obtained using the cdf and is equal to 0.6490. The following figures show the marginal pmf of X and Y.



(a)



(b)

Independent discrete random variable

If the events $X = x$ and $Y = y$ are statistically independent (or X and Y are statistically independent), the contidional pmf is:

$$p_{X|Y}(x|y) = p_X(x) \quad \text{and} \quad p_{Y|X}(y|x) = p_Y(y)$$

and the joint pmf is:

$$p_{X,Y}(x, y) = p_X(x)p_Y(y)$$

Joint probability distribution of continuous variables

Joint pdf and cdf for continuos X and Y variables

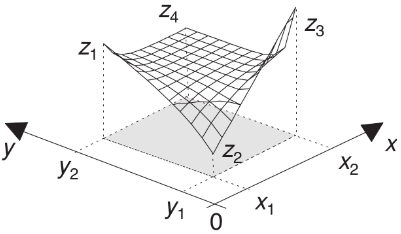
If two continuos random variables X and Y ocurr simultaneously or are related somehow, the joint probability distribution is described by the **joint probability density function** (pdf), $f_{X,Y}(x,y)$. The probability over a region of interest $\{(x_1,x_2),(y_1,y_2)\}$, where $x_1 < x_2$ and $y_1 < y_2$, is defined as:

$$Pr[(x_1 \leq X \leq x_2) \cap (y_1 \leq Y \leq y_2)] = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{X,Y}(x,y)dydx$$

As shown in the figure, this integral represent the volume below the surface represented by the joint pdf over the region of interest. According to the probability axioms, the $f_{X,Y}(x,y)$ has the following properties:

$$f_{X,Y}(x,y) \geq 0$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = 1$$



The **joint cumulative distribution function** (cdf), $F_{X,Y}(x,y)$, is defined as:

$$F_{X,Y}(x,y) \equiv Pr[(-\infty \leq X \leq x) \cap (\infty \leq Y \leq y)] = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y)dydx$$

Joint probability distribution of continuous variables

Joint pdf and cdf for continuos X and Y variables

These concepts can be applied to n random variables X defined on the same probability space. So $(X_1, X_2, \cdots, X_n)$ is an **n-dimensional continuous random variable** if $f_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) \geq 0$. Accordingly:

$$
\begin{aligned}
&F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) \\
&\equiv Pr[(-\infty \leq X_1 \leq x_1) \cap (\infty \leq X_2 \leq x_2) \cap \cdots \cap (\infty \leq X_n \leq x_n)] \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n
\end{aligned}
$$

where $(x_1, x_2, \cdots, x_n)$ is a n-tuple of points in the sample space. Recalling from the univariate case, the pdf and the cdf are related for the bivariate case as:

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$$

For n random variables:

$$f_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n) = \frac{\partial^n}{\partial x_1 \partial x_2 \cdots \partial x_n} F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n)$$

Joint probability distribution of continuous variables

Storm intensity and duration

An storm occurring in place on earth is characterized its intensity and its duration, where the intensity is the average amount of rain fell. Suppose that intensity and duration are two continuous random variables, Y and X, respectively. It is known that the cdf of X and Y are represented by:

$$F_X(x) = 1 - e^{-ax},\ x \geq 0, a > 0; \quad F_Y(y) = 1 - e^{-by},\ y \geq 0, b > 0;$$

where a and b are parameters of the cdfs. Accordingly, it is assumed that the joint cdf is given by the exponential bivariate distribution:

$$F_{X,Y}(x, y) = 1 - e^{-ax} - e^{-by} + e^{-ax-by-cxy}$$

where c is a parameter of the joint cdf that describes the joint variability (or correlation) of X and Y . According to the marginal cdfs, a and b are > 0, so that, one might be interested to known the possible values of c. To search for the lower bound of c, it is known by definition that $F_{X,Y}(x, y) \leq F_X(x)$, because the joint $Pr[X \leq x, Y \leq y]$ can not exceed $Pr[X \leq x]$ independently of the value taken by Y. The same apply for Y. Accordingly:

$$F_{X,Y}(x, y) = 1 - e^{-ax} - e^{-by} + e^{-ax-by-cxy} \leq F_X(x) = 1 - e^{-ax}$$

resolving:

$$-e^{-by} - e^{-ax-by-cxy} \leq 0 \quad e^{-ax-by-cxy} \leq e^{-by} \quad -x(a + cy) \leq 0$$

Regarding that X and Y are non negative variables, the inequality $-x(a + cy) \leq 0$ holds if and only if $(a + cy) \geq 0$.

Joint probability distribution of continuous variables

Storm intensity and duration

To determine the upper bound of c we need to get the joint pdf. So that, first we get the partial derivative with respect to x and then with respect to y:

$$\frac{\partial F}{\partial x} = \frac{\partial \left(1 - e^{-ax} - e^{-by} + e^{-ax-by-cxy}\right)}{\partial x} = ae^{-ax} - (a+cy)e^{-ax-by-cxy}$$

$$\begin{aligned} f_{X,Y}(x,y) = \frac{\partial^2 F}{\partial x \partial y} &= \frac{\partial \left(ae^{-ax} - (a+cy)e^{-ax-by-cxy}\right)}{\partial y} \\ &= \left[(a+cy)(b+cx) - c\right] e^{-ax-by-cxy} \end{aligned}$$

Accordingly, the for $x = y = 0$, $f_{X,Y}(0,0) = ab - c$. Since the joint pdf is a non-negative function, the inequality $ab - c \geq 0$ must hold, so the upper bound of c is $c \leq ab$. Summing up, the bivariate exponential distribution is defined for $0 \leq c \leq ab$.

Joint probability distribution of continuous variables

Conditional probability density function

The **conditional density function** of Y given X is written as:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

From where,

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) = f_{X|Y}(x|y)f_Y(y)$$

Storm intensity and duration

From the example above, it is known that $a > 0$, $b > 0$ and $0 \leq c \leq 1$ are three parameters estimated based on rainfall data. For weather station, the following values were estimated: $a = 0.05$ h^{-1}, $b = 1.2$ h/mm and $c = 0.06$ mm^{-1}. Engineers are planning to design a drainage system, so they need to estimate the probability that an storm lasting $X = 6$ hours will exceed an average intensity of $Y = 2$ mm/h. The conditional pdf of the storm intensity for a given duration is:

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y(x,y)}}{f_X(x)} = \frac{\left[(a+cy)(b+cx)-c\right]e^{-ax-by-cxy}}{ae^{-ax}} \\ &= a^{-1}\left[(a+cy)(b+cx)-c\right]e^{-y(b+cx)} \end{aligned}$$

Storm intensity and duration

The conditional cdf is (making $y \equiv u$):

$$F_{Y|X}(y|x) = \int_0^y a^{-1}\left[(a+cu)(b+cx) - c\right] e^{-u(b+cx)} du$$
$$= a^{-1}(b+cx)\int_0^y (a+cu)e^{-u(b+cx)} du - a^{-1}c \int_0^y e^{-u(b+cx)} du$$
$$= 1 - \frac{a+cy}{a} e^{-y(b+cx)}$$

Evaluating the following:

$$Pr[Y > 2|X = 6] = 1 - F_{Y|X}(2|6) = 1 - 1 + \frac{0.05 + 0.06x2}{0.05} e^{-2(1.2+0.06x6)} = 0.15$$

Joint probability distribution of continuous variables

Independent continuous random variable

If the events $X = x$ and $Y = y$ are stochastically independent, then $f_{X|Y}(x, y) = f_X(x)$ and $f_{Y|X}(y, x) = f_Y(y)$, therefore $f_{X,Y}(x, y) = f_X(x)f_Y(y)$. Note that the assumption of independence simplifies the application of probability to engineering problems.

Storm intensity and duration

If the joint variability of storm duration X and storm intensity Y, is neglected, the joint pdf is given by:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = ae^{-ax}be^{-by} = ab\ e^{-ax-by}$$

The same expression is valid for $c = 0$.

Marginal probability density function

The extension of the total probability theorem gives:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dy = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x)f_X(x)dx = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx$$

Storm intensity and duration

The joint pdf for the storm duration X and the storm intensity Y is $f_{X,Y}(x,y) = [(a+cy)(b+cx)-c]\,e^{-ax-by-cxy}$. From the definition of marginal probability density function:

$$f_X(x) = \int_0^\infty f_{X,Y}(x,y)dy = \int_0^\infty [(a+cy)(b+cx)-c]\,e^{-ax-by-cxy}dy = a\ e^{-ax}$$

$$f_Y(y) = \int_0^\infty f_{X,Y}(x,y)dx = \int_0^\infty [(a+cy)(b+cx)-c]\,e^{-ax-by-cxy}dx = b\ e^{-by}$$

Properties of multiple variables

Covariance and correlation

The expectation operator (E) introduced in below for a single variable can be used on two or more random variables. For instance, for the linear combination of two random variables X_1 and X_2, $E[aX_1 + bX_2] = aE[X_1] + bE[X_2]$, where a and b are constants. Accordingly, the variance of $aX_1 + bX_2$ is:

$$Var[aX_1 + bX_2] = a^2 Var[X_1] + b^2 Var[X_2] - 2abCov[X_1, X_2]$$

where the operator Var is the variance operator and Cov is the **covariance** operator. The **covariance** of X_1 and X_2 is calculated as:

$$Cov[X_1, X_2] = E[(X_1 - E[X_1])(X_2 - E[X_2])] = E[X_1 X_2] - E[X_1]E[X_2]$$

where $E[X_1 X_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X_1,X_2}(x_1, x_2) d_{x_1} d_{x_2}$. If X_1 and X_2 are independent:

$$\begin{aligned}
E[X_1 X_2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X_1,X_2}(x_1, x_2) d_{x_1} d_{x_2} \\
&= \left[\int_{-\infty}^{\infty} x_1 f_{X_1}(x_1) dx_1\right] \left[\int_{-\infty}^{\infty} x_2 f_{X_2}(x_2) dx_2\right] \\
&= E[X_1]E[X_2]
\end{aligned}$$

therefore $Cov[X_1, X_2] = 0$ and $Var[aX_1 + bX_2] = a^2 Var[X_1] + b^2 Var[X_2]$.

Properties of multiple variables

Storm intensity and duration

The joint pdf of the independent random variables, storm duration X and storm intensity Y, is $f_{X,Y}(x,y) = f_X(x)f_Y(y) = a\ e^{-ax} b\ e^{-by} = ab\ e^{-ax-by}$, where $\mu_X = a^{-1}$ and $\mu_Y = b^{-1}$. To estimate $Cov[X,Y]$, it is needed to estimated:

$$E[XY] = \int_0^\infty \int_0^\infty xyf_{X,Y}(x,y)dxdy = \int_0^\infty \int_0^\infty xy\ ab\ e^{-ax-by}dxdy = \frac{1}{ab}$$

The $Cov[X,Y]$ is:

$$Cov[X,Y] = E[XY] - E[X]E[Y] = \frac{1}{ab} - \frac{1}{a}\frac{1}{b} = 0$$

This confirm that there is no covariance between X and Y.

Coefficient of linear correlation

Note that $Cov[X_1, X_2]$ is positive and large when both variables are large or small with respect to their means.In contrast, $Cov[X_1, X_2]$ can be negative and large when both variables are equidistand (e.g. one small and the other large). Note that covariance is a mesuare of the linear relationship between variables; no linearly related variables yield no covariance. The linear relationship between X_1 and X_2 is defined as the **coefficient of linear correlation (ρ)**:

$$\rho = \frac{Cov[X_1, X_2]}{\sigma_{X_1}\sigma_{X_2}}$$

where $-1 \leq \rho \leq 1$.

Properties of multiple variables

Joint moment-generating function

The **Joint moment-generating function** for multiple random variables $X_1, X_2, \cdots, X_k$ is:

$$M_{X_1,X_2,\cdots,X_k}(t_1, t_2, \cdots, t_k) = E\left[e^{\left(\sum_{i=1}^{k} t_i X_i\right)}\right]$$

where the rth moment of X_i can be computed by diferentiating the joint moment-generating function r times with respect to t_i and then evaluating the derivative for $t_i = 0$. For two statistically independent random variables, X_1 and X_2, the Joint moment-generating function is:

$$M_{X_1,X_2}(t_1, t_2) = E[e^{t_1 X_1 + t_2 X_2}] = E[e^{t_1 X_1}]E[e^{t_2 X_2}] = M_{X_1}(t_1)M_{X_2}(t_2)$$

This means that if the joint mgf of two variables equal the product of individual mgfs, the two variables are statistically independent.

Storm intensity and duration

The joint pdf of the independent random variables, storm duration X and storm intensity Y, is $f_{X,Y}(x, y) = ab\ e^{-ax-by}$, where $\mu_X = a^{-1}$ and $\mu_Y = b^{-1}$. The joint mgf of X and Y is:

$$M_{X,Y}(t_1, t_2) = E\left[e^{t_1 X + t_2 Y}\right] = E\left[e^{t_1 X}\right] E\left[e^{t_2 Y}\right] = \frac{ab}{(a - t_1)(b - t_2)}$$

Probability distributions

Generalities

- From the data sample representation is possible to obtain the **empirical distribution**.
- The **probability distributions** are mathematical functions that satisfy the probability axioms.
- These distributions or functions are appropriate to describe analytically the random behaviour of certain variables and are interpreted as probability models.
- Accordingly, the probability distributions establish the probability of occurrence of a value taken by the random variable.
- A random variable can be represented by multiple distributions, however the ´closest one´ to the empirical distribution must be chosen.
- In hydrology, there are certain distributions that adjust the best to some hydrological variables according to decades of experience.
- To fit a probability distribution to a random variable upon a random variable sample is needed to adjust the distribution **parameters**.
- The more parameters the function has, the most flexible is the distribution to fit the empirical distribution. However the more parameters, the less degrees of freedom.
- Probability distributions can according to the random variable type:
  - Discrete probability distributions: Used for discrete random variables
  - Continuous probability distributions: Used for continuous random variables

Discrete probability distributions

The **binomial** distribution

This distribution is used to define the probabilities of discrete random variables. For a sample of size n an instance of the random variable only take two values (0 or 1, success or failure). The n observations or experiments are mutually exclusive and collectively exhaustive, and the probability of occurrence of either value is constant. Each observation or experiment follow a **Bernoulli distribution** where the probability of success is p and the probability of failure is $1-p$. The set of n observations or experiments is known as the experiments of Bernoulli from where a random variable X, which represents the observation with success in n observations, follows a **Binomial distribution**:

$$p_X(x) = Pr[X = x; n, p] = B(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

where n and p are parameters of the distribution, $x = 0, 1, \cdots, n$, $0 \leq p \leq 1$ and $\binom{n}{k} = \frac{n!}{x!(n-x)!}$; this term represents the possible combinations to obtain x success from n. The cmf is:

$$F_X(x) = Pr[X \leq x] = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$$

For this distribution, the mean is $E[X] = np$ and the variance is $Var[X] = np(1-p)$.

Discrete probability distributions

Flooding of a road

The probability that a road is flooded once in a give year is $p = 0.1$. Compute the probability that the road will be flooded at least once during a five-year period. So the $Pr[X \geq 1] = Pr[X = 1; 5, 0.1] + Pr[X = 2; 5, 0.1] + Pr[X = 3; 5, 0.1] + Pr[X = 4; 5, 0.1] + Pr[X = 5; 5, 0.1] = 0.3281 + 0.0729 + 0.0081 + 0.0005 + 0.00001 = 0.4095$. An alternative to estimate this probability is to determine the probability of no flooded during the period ($Pr[X < 1] = 1 - Pr[X \geq 1]$). So $Pr[X \geq 1] = 1 - Pr[X = 0; 5, 0.1] = 1 - 0.5905 = 0.4095$. Note that in reality, the values of p can change year to year so the application of this model can be inaccurate.

The geometric distribution

This distribution characterizes the probability of a random variable X that represents the Bernoulli's trials upon one get a success. So:

$$P_X(x) = Pr[X = x; p] = (1 - p)^{x-1}p$$

where p ($0 < p \leq 1$) is the probability of success in the Bernoulli's trials, and $X = 0, 1, 2, \cdots, x - 1$ failures before the first success. A converging geometric series $\sum_{x=1}^{\infty}(1 - p)^x = \frac{1-p}{p}$. To optain the mean, one can derive both sides of the geometric series with respect to p and multiply by $-p$, which is equivalent to $E[X] = \sum_{x=1}^{\infty} x(1 - p)^{x-1}p = \frac{1}{p}$. To obtain $Var(X) = E[X(X - 1)] = E[X^2] - E[X] = \sum_{x=1}^{\infty} x(x - 1)(1 - p)^{x-1}p =$

The **geometric** distribution

To obtain $Var(X)$ one use the second factorial moment as $E[X(X-1)] = E[X^2] - E[X] = \sum_{x=1}^{\infty} x(x-1)(1-p)^{x-1}p$. Getting the second derivative of the geometric series with respect to p and multiplying both sides by $p(1-p)$, one get that $E[X^2] - E[X] = \frac{2(1-p)}{p^2}$. Regarding this and that $Var(X) = E[X^2] - (E[X])^2 = \frac{2(1-p)}{p^2} + E[X] - (E[X])^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$

Continuous probability distributions

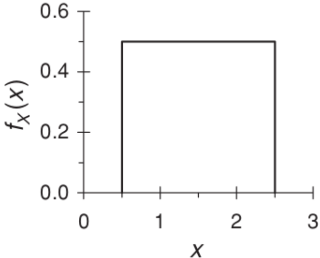
Continuous probability distributions are applicable when the random variable is continuous; take any value in its real domain, e.g. river flow, air temperature.

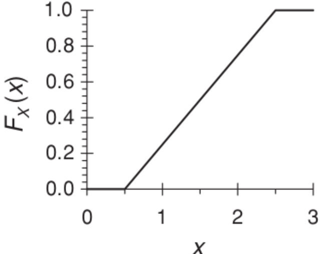
Uniform distribution

It is the simplest distribution, and as it is indicated by its name, the pdf is constant over a defined interval $a \leq x \leq b$. The pdf is:

$$f_X(x) = \frac{1}{b-a}, \text{ for } a \leq x \leq$$
$$= 0, \text{ otherwise}$$

This distribution is also knows as the rectangular distribution. Note that all values between a and b are equally probable and the area under the pdf is equal to 1. The probability that X fall in (c, d) is $Pr[c < X < d] = \frac{d-c}{b-a}$. The unit uniform distribution is obtained when $a = 0$ and $b = 1$ ans it is commonly used to generate random variables in simulations.





Continuous probability distributions

Uniform distribution

- Mean

$$E[X] = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{b+a}{2}$$

- Variance

$$\begin{aligned} Var[X] = E[X^2] - (E[X])^2 &= \int_a^b \frac{x^2}{b-a} dx - (E[X])^2 \\ &= \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b - \frac{(b+a)^2}{4} \\ &= \frac{b^3 - a^3}{3(b-a)} - \frac{(b+a)^2}{4} = \frac{b^2 + ba + a^2}{3} - \frac{(b+a)^2}{4} \\ &= \frac{(b-a)^2}{12} \end{aligned}$$

- Moment-generating function

$$M_X(t) = E\left[e^{tX}\right] = \int_a^b \frac{e^{tx}}{b-a} dx = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

Exponential distribution

From the **Poisson process** discussed previously, if we denote by random variable T the time to the first arrival, then the probability that T exceeds some value t is equal to the probability that no events occur in that time interval of length t. While former probability is 1 - $F_T(t)$, the latter probability $p_X(0)$ is zero, which is the probability that a Poisson random variable X with parameter λt. Replacing:

$$1 - F_T(t) = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t} \quad t \geq 0$$

Therefore:

$$F_T(t) = 1 - e^{-\lambda t} \quad t \geq 0$$

and the $f_t(t)$ is:

$$f_T(t) = \frac{dF_T(t)}{dt} = \lambda e^{-\lambda t} \quad t \geq 0$$

This equation define the **exponential distribution**, which describes the time to the first occurence of a Poisson event. Note that $e^{-\lambda t}$ is the probability of no events in any interval of time of length t, whether or not it begins at time 0. In short, the interarrival times of a **Poisson process** are independent and exponentially distributed.

- Mean

$$E[T] = \int_0^\infty t\lambda e^{-\lambda t} dt$$

  making $u = \lambda t$:

$$E[T] = \frac{1}{\lambda} \int_0^\infty u e^{-u} du = \frac{1}{\lambda} \left[e^{-u}(-u - 1) \right]_0^\infty = \frac{1}{\lambda}$$

Continuous probability distributions

Exponential distribution

Note that if λ is the rate at which events occur in a Poisson process, $\frac{1}{\lambda}$ is the average time between events.

- Variance

$$
\begin{aligned}
Var[T] = E[T^2] - (E[T])^2 &= \int_0^\infty t^2 \lambda e^{-\lambda t} dt - (E[X])^2 \\
&= \frac{1}{\lambda^2} \int_0^\infty u^2 e^{-u} du - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}
\end{aligned}
$$

It easy to show that the coefficient of variation $V_T = \frac{\sqrt{Var[T]}}{E[T]} = 1$.

- Moment-generating function

$$
M_X(t) = E\left[e^{tX}\right] = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}, \quad \text{for } t < \lambda
$$

Floods affecting construction

Engineers are concerned of the occurrence of a flood exceeding 100 m^3 s^{-1} that can seriously affect infrastructure. If this streamflow is exceeded once every five years on average based on historical flow records, what is the change that the infrastructure construction which is scheduled to last 14 months can proceed without interruptions or detrimental effects? Suppose that flow exceeding this magnitude are independent and identically distributed events. The sample mean is $\bar{t} = 5$ years, thus $\bar{\lambda} = \frac{1}{5}$. From the exponential distribution, $Pr\left[X \geq \frac{14}{12}\right] = e^{-1/5.14/12} = 0.79$. Accordingly, the risk is thus 1-0.79 $= 0.21$, which is quite high. The solution would be to shorten the period of construction.

Exponential distribution

- Memoryless property The Poisson process is often said to be memoryless meaning that future behaviour is independent of its present or past behaviour. This memoryless trait of the Poisson arrivals and of the exponential distribution is best understood by determining the conditional distribution of T given that $T > t_0$, that is, the distribution of the time between arrivals given that no arrivals occurred before t_0:

$$F_{T|[T>t_0]}(t) = P[T \leq t | T > t_0] = \frac{P\left[(T \leq t) \cap (T > t_0)\right]}{P[T > t_0]}$$

For $t < t_0$, the numerator is zero; for $t \geq t_0$ it is simply equal to $P[t_0 < T \leq t]$. Thus:

$$F_{T|[T>t_0]}(t) = \frac{F_T(t) - F_T(t_0)}{1 - F_T(t_0)} = \frac{(1 - e^{-\lambda t}) - (1 - e^{-\lambda t_0})}{e^{-\lambda t_0}}$$

$$= \frac{e^{-\lambda t_0} - e^{-\lambda t}}{e^{-\lambda t_0}} = 1 - e^{-\lambda(t - t_0)} \quad t \geq t_0$$

$$f_{T|[T>t_0]}(t) = \lambda e^{-\lambda(t-t_0)} \quad t \geq t_0$$

If $\tau = t - t_0$,

$$f_{T|[T>t_0]}(t_0 + \tau) = \lambda e^{-\lambda \tau} \quad \tau \geq 0$$

This means that failure to observe an event up to t_0 does not alter one's prediction of the length of time (from t_0) before an event will occur. The future is not influenced by the past if events are Poisson arrivals. An implication is that any choice of the time origin is satisfactory for the Poisson process.

Continuous probability distributions

Triangular distribution

This distribution is defined in the range of $a \leq X \leq b$ and the pdf is:

$$f_X(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{if } a \leq x \leq c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{if } c < x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Note that c correspond to the highest triangle vertex and it is the mode of pdf. The cdf is thus:

$$F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{(x-a)^2}{(b-a)(c-a)} & \text{if } a \leq x \leq c \\ 1 - \frac{(b-x)^2}{(b-a)(b-c)} & \text{if } c < x \leq b \\ 1 & \text{otherwise} \end{cases}$$

- Mean
$$E[X] = \int_a^b x f_X(x) dx = \frac{a+b+c}{3}$$
- Varianza
$$Var[X] = E[X^2] - (E[X])^2 = \int_a^b x^2 f_X(x) dx - (E[X])^2 = \frac{a^2+b^2+c^2-ab-ac-bc}{18}$$

This distribution is used in hydrology as an alternative to the **uniform distribution** to describe the sample space and the uncertainty of parameters in hydrological modelling.

Continuous probability distributions

The Gamma distribution

As in the case with discrete trials, it is also of interest to ask for the distribution of the time X_k to the kth arrival of a **Poisson process**. Now, the times between arrivals, T_i, $i = 1, 2, \cdots, k$, are independent and have an exponential distribution with common parameter λ. X_k is the sum $T_1 + T_2 + \cdots + T_k$. Its distribution follows from repeated application of the convolution integral $f_Z(z) = \int_{\infty}^{\infty} f_X(z-y) f_Y(y) dy$, where $Z = X + Y$ and X and Y are independent. For any $k = 1, 2, 3, \cdots$:

$$f_{X_k}(x) = \frac{\lambda(\lambda x)^{k-1} e^{-\lambda x}}{(k-1)!} \quad x \geq 0$$

Thus, we say that X follow a Gamma distribution with parameters k and λ. Regarding that X can be considered as the sum of k independent exponentially distributed random variables, then:

- Mean

$$E[X] = \int_0^{\infty} x f_{X_k}(x) dx = \frac{k}{\lambda}$$

- Variance

$$Var[X] = E[X^2] - (E[X])^2 = \int_0^{\infty} x^2 f_{X_k}(x) dx - (E[X])^2 = \frac{k}{\lambda^2}$$

The denominator of the pdf $(k-1)!$ is the product of the first $(k-1)$ natural numbers and can be written as the **standard gamma function** $\Gamma(k)$.

Continuous probability distributions

The Gamma distribution

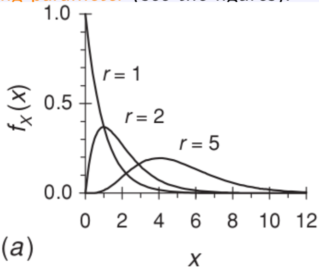
$\Gamma(k)$ is applicable also to non integer values of k and is written as:

$$\Gamma(k) = \begin{cases} \int_0^\infty u^{k-1} e^{-u} du & \text{for } k > 0 \\ 0 & \text{otherwise} \end{cases}$$

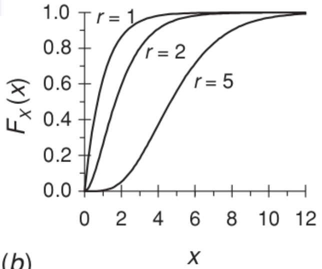
Integrating by parts, it can be shown that $\Gamma(k+1) = k\Gamma(k)$ for any $k > 0$. Also, when $k = 1$ $\Gamma(1) = 1$ and when $k = 1/2$ $\Gamma(1/2) = \sqrt{\pi}$. The **standard gamma** pdf, can be written as:

$$f_X(x) = \begin{cases} \frac{\lambda(\lambda x)^{k-1} e^{-\lambda x}}{\Gamma(k)} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

for $k > 0$ and $\lambda > 0$. Note that $\Gamma(k)$ arise here as constant to normalize the function. k is denominated as the shape parameter as it defines the shape of the pdf and λ the scaling parameter (see the figures).



(a)



(b)

The Gamma distribution

The $\Gamma(k)$ function is widely used as the incomplete gamma function as:

$$\Gamma(k,x) = \begin{cases} \int_0^x u^{k-1} e^{-u} du & \text{for } k > 0 \\ 0 & \text{otherwise} \end{cases}$$

This can be used to evaluate the gamma cdf:

$$F_X(x) = \begin{cases} \int_0^x f_X(X) dx = \frac{\Gamma(k,\lambda x)}{\Gamma(k)} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

It has been found that empirical distributions of many natural and non-natural processes that take positive values closely resemble gamma distribution. This distribution has been implemented to describe phenomena such as maximum stream flows and the depth of monthly precipitation. In statistics, the gamma distribution is important because the **chi-squared distribution** is a particular form of the gamma with $k = \nu/2$, where ν are the degrees of fredom, and $\lambda = 1/2$. The cdf of the chi-squared distribution is:

$$F(\chi^2) = \frac{1}{2} \int_0^{\chi^2} \frac{(t/2)^{(\nu/2)-1} e^{-t/2}}{\Gamma(\nu/2)} dt$$

where t is a dummy variable.

Maximum flows

Based on a histogram of data of maximum annual river flows in the Weldon River, USA, between 1930 and 1960, it was assume that the data followed a gamma pdf. The parameters were estimated and equal to $k = 1.727$ and $\lambda = 0.00672$ cfs^{-1}. According to these values, $\mu_X = \frac{k}{\lambda} = \frac{1.727}{0.00672} = 256.7$ cfs, $\sigma_X = \frac{\sqrt{k}}{\lambda} = 190$ cfs. Also, the probability that the maximum flow is less than 400 cfs in any year is $F_X(400) = \frac{\Gamma(1.727, \lambda 400}{\Gamma(1.727)} = \frac{0.71}{0.914} = 0.78$. No that the values to estimate the numerator and denominator are taken from tables of the gamma function.

The Beta distribution

The **Beta distribution** is uses to model a random variable whose values go from 0 to 1, and thus it is important in decision methods. The beta pdf is given by:

$$f_X(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where α and β are parameters that $\alpha > 0$ and $\beta > 0$, and:

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

From this two equations, the nth moment is:

$$E[X^n] = \frac{B(\alpha+n,\beta)}{B(\alpha,\beta)}$$

Considering that $\Gamma(k+1) = k\Gamma(k)$) and using this equation, we have:

$$E[X] = \frac{\alpha}{\alpha+\beta}$$

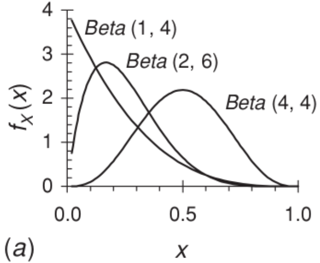
Similarly, considering that $\Gamma(k+2) = (k+1)\Gamma(k+1) = (k+1)k\Gamma(k)$ and simplifying,

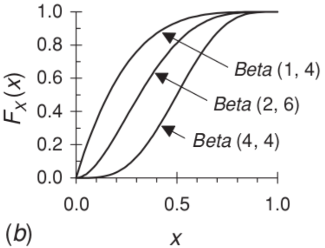
$$Var[X] = E[X^2] - (E[X])^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Continuous probability distributions

The Beta distribution

The following figures show the pdf and the cdf for different combinations of α and β.





Note that when $\alpha = \beta$ the pdf is symmetrical and when $\alpha = \beta = 1$ is equivalent to the uniform distribution with $a = 0$ and $b = 1$.

Continuous probability distributions

The Weibull distribution

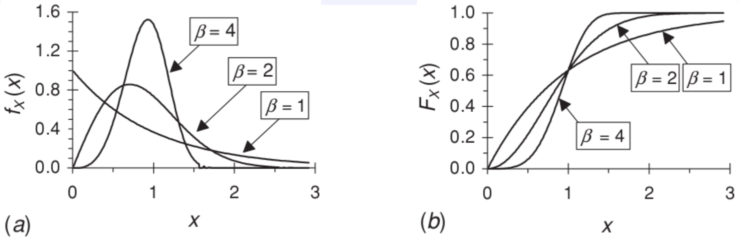
The **Weibull distribution** approximate closely to many natural phenomena. It has been used to model, for instance, the time to failure of electrical and mechanical systems. The pdf is defined as:

$$f_X(x) = \begin{cases} \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-\left(\frac{x}{\lambda}\right)^{\beta}}, & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

And the cdf takes the form:

$$F_X(x) = \begin{cases} f_X(x) = \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-\left(\frac{x}{\lambda}\right)^{\beta}}, & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda > 0$. The figures below show the pdf and cdf for different values of β and $\lambda = 1$.



Note than from the definition of the pdf and cdf, if a random variable $X \sim$ Weibull (β, λ), then $Y = (X/\lambda)^{\beta} \sim$ exponential $(\lambda = 1)$, $f_Y(y) = e^{-y}$.

Continuous probability distributions

The **Weibull** distribution

- mean

$$E[X] = \lambda\Gamma\left(1 + \frac{1}{\beta}\right)$$

- variance

$$Var[X] = E[X^2] - (E[X])^2 = \lambda^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2\right]$$

Note that for $\beta = 1$ the Weibull pdf becomes the exponential pdf.

Estimation of low flows

From a ten-years daily average streamflow time series in cfs at the River Pang at Pangbourne 10-days averages are estimated. From this new aggregated time series, the minimum annual 10-days average flows are estimated. The ranked values are: 13.4 25.7 32.2 35.9 40.0 40.0 40.4 50.7 58.2 71.4. Regarding that the Weibull distribution is associated with the extreme value theory, the distribution can be applied to minima. Estimating the mean and the variance from the data and using the equations yield by the method of moments (see the equations above), one can resolve this two equations iteratively for the shape parameter β and for the scale parameter λ. Alternatively, the methods of likelihood can be used iteratively. An alternative procedure is to used the least-squares procedure from the cdf. So from the cdf:

$$\ln(x_i) - \ln(\lambda) = \frac{[\ln[-\ln(1 - F_X(x_i))]]}{\beta}$$

Continuous probability distributions

Estimation of low flows

Applying the plotting position, which is the probability at which x_i should be plotted, we have $F_X(x_i) = \frac{(i-0.35)}{n}$, and letting $z_i = \ln(x_i)$ and $y_i = \ln\left[-\ln\left(1 - F_X(x_i)\right)\right]$, we obtain:

$$z_i = \frac{y_i}{\beta} + \ln(\lambda) + \epsilon$$

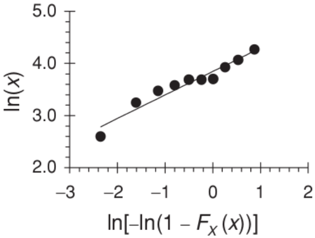
where i is the rank of the data in ascending order, n is the number of data values and ϵ is an error term in the regression. Applying the least-square fit to this equation for $n = 10$:

$$\hat{\beta} = \frac{\sum_{i=1}^{n=10} (y_i - \bar{y})^2}{\sum_{i=1}^{n=10} (y_i - \bar{y})(z_i - \bar{z})} = 2.59$$

and

$$\hat{\lambda} = e^{\left(\bar{z} - \bar{y}/\hat{\beta}\right)} = 44.85\, m^3/s$$

where $\bar{z} = \frac{1}{10}\sum_{i=1}^{n=10} z_i$ and $\bar{y} = \frac{1}{10}\sum_{i=1}^{n=10} y_i$. This linear relationship between z and y represent the least-squares fit of the data to the two-parameter Weibull distribution (see figure).



The approximate lower 99% confidence limit for β is $\frac{\hat{\beta}}{2n}\chi^2_{2n,0.99} = \frac{2.59}{20}8.26 = 1.07$. Also, quantile estimates can be obtained using the estimates of the Weibull parameters and the preceding equations. For example the annual minimum 10-day flow with a return period of 10 yrs is:

$$y_{10} = \ln\left[-\ln\left(1 - \frac{1}{10}\right)\right] = -2.25$$

$$x_{10} = e^{\left[\left(\frac{y_{10}}{\hat{\beta}}\right) + \ln(\hat{\lambda})\right]} = 18.8\, m^3/s$$

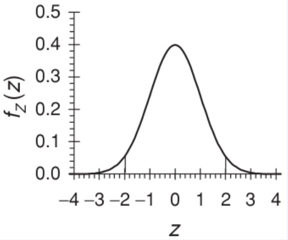
The Normal distribution

The **Normal or Gauss distribution** originally emerged in the study of experimental errors. This errors are differences between observations recorded under unchanged similar experimental conditions. In telecommunications, errors are used to be known as noise, which is, in general, the difference between the true estate and the observation. The **Normal distribution** is thus ideal to represent such error when they are of an additive nature. The pdf of the **Normal distribution** of a random variable X is:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]}, \text{ for } -\infty < x < \infty$$

This pdf is especified by two parameters: the mean or location parameter (μ) and the standard deviation or scale parameter (σ), computed from the population. In practices, it is common to use the standardized curve with the transformation of X into Z as $Z = \frac{X-\mu}{\sigma}$, so the pdf is:

$$f_Z(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}z^2\right)}$$



When $\mu = 0$ and $\sigma = 1$, the pdf is:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}z^2\right)}$$

The Normal distribution

Note that the cdf can only be calculated using numerical methods and it is:

$$F_X(x) = Pr[X \leq x] = Pr\left[Z \leq \frac{X-\mu}{\sigma}\right] = Pr[Z \leq z]$$

$$= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{u^2}{2}\right)} du, \text{ for } -\infty < z < \infty$$

Because of the symmetry of the pdf, $f_Z(-z) = 1 - f_Z(z)$. Below, some properties of the normal distribution:

- A linear transformation $Y = bX + a$ of a random variable X with $N \sim (\mu, \sigma^2)$ makes Y an $N \sim (b\mu + a, b^2\sigma^2)$ random variable.
- If X_i $i = 1, 2, \cdots, n$ are independent and identically distributed random variables from a population with mean μ and standard deviation σ, then the random variable $\bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n}$ that is the sample mean, from a random sample of size n from the sample population, tends to have an $N \sim (\mu, \sigma^2/n)$ distribution as $n \to \infty$. This important results is called the **central limit theorem**. The theorem states that even if the distribution of the random variable X_i is nor normal, the so called sampling distribution of its mean will tend to normality asymptotically.
- If X and Y are independent random variables normally distributed with μ_X and μ_Y, and σ_X^2 and σ_Y^2, respectively, the expression $Z = X + Y$ is normally distributed too with $\mu_Z = \mu_X + \mu_Y$ and $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$. This is also applicable for $Z = X - Y$ with $\mu_Z = \mu_X - \mu_Y$ and $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$. If X and Y are not statistically independent but correlated, Z is still normally distributed bu $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y$, where ρ is the linear correlation between X and Y.

Streamflows

Annual average discharges in river X are distributed following a $N \sim (400, 50^2)$.Annual average discharges of a tributary Y of river X are distributed as $N \sim (145, 76^2)$. Downstream of the mouth of river Y, there is an outflow D to supply water to agricultural fields. Downstream of D, annula avearge discharges in river X are distributed as $N \sim (425, 91.5^2)$. Assuming that X, Y and D are independent, the pdf of annual average discharges in D is normal with $\mu_D = 300 + 145 - 425 = 20$ and $\sigma_D^2 = 91.5^2 - 76^2 - 50^2 = 9.81^2$. The probability that $D > 30$ can be estimated as $z = \frac{30-20}{9.81} = 1.0194$ and $F_Z(z = 1.0194) = 0.846$. So the $Pr[D > 30] = 1 - 0.846 = 0.154$. If X, Y and D are not independent and $\rho_{X,Y} = 0.15$ and $\rho_{X,D} = -0.35$ (under the assumption that that the lower the streamflow in X the greater the flow derived by D), $\mu_D = 20$ and $91.5^2 = 97.0261^2 + \sigma_D^2 - 2(-0.35)(97.0361)\sigma_D$, from where $\sigma_D^2 = 23.49^2$. For this case, the $Pr[D > 30] = 0.3352$.

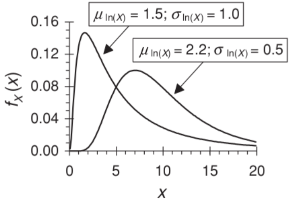
Continuous probability distributions

The Lognormal distribution

We know that the addition of large number of small random effects will tend to make the distribution of the aggregate approach normal. In contrast, if the phenomena arises from the multiplicative effect of a large number of uncorrelated factors, the distribution tend to be **Lognormal**, that is, the logarithm of a random variable X becomes normally distributed. Phenomena such as the particle size of sediment samples are Lognormal, and the interarrival times of earthquakes. As this distribution of defined for non-negative values is used to describe daily average streamflows, maximum streamflows, and daily, monthly and annual rainfall. If X is a positive random variable, then $Y = \ln(X)$. Accordingly, if Y has $N \sim (\mu_Y, \sigma_{Y^2})$, X has a Lognormal distribution $LN \sim (\mu_{\ln(X)}, \sigma^2_{\ln(X)})$. So the pdf is:

$$f_X(x) = \frac{1}{x\sigma_{\ln(X)}\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left[\frac{\ln(X)-\mu_{\ln(X)}}{\sigma_{\ln(X)}}\right]^2\right]}, \text{ for } 0 \leq x < \infty$$

See the figure below with plot of the Lognormal pdf.



The expectation operator can be written as:

$$E\left[X^r\right] = \int_0^\infty x^r f_X(x)dx$$

$$= \frac{1}{\sigma_{\ln(X)}\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ry} e^{\left[-\frac{1}{2}\left[\frac{y-\mu_{\ln(X)}}{\sigma_{\ln(X)}}\right]^2\right]} dy$$

The Lognormal distribution

From where $E[X^r] = E\left[e^{Yr}\right]; Y \sim N\left(\mu_{\ln(X)]}, \sigma^2_{\ln(X)}\right)$, so $E[X^r] = M_Y(r) = e^{r\mu_{\ln(X)} + \frac{1}{2}r^2\sigma^2_{\ln(X)}}$. Thus , the mean is:

$$\mu_X = E[X] = e^{\left[\mu_{\ln(X)} + \frac{1}{2}\sigma^2_{\ln(X)}\right]}$$

and the variance is:

$$\sigma^2_X = E[X^2] - \mu^2_X = e^{2\left[\mu_{\ln(X)} + \sigma^2_{\ln(X)}\right]} - e^{2\left[\mu_{\ln(X)} + \frac{1}{2}\sigma^2_{\ln(X)}\right]} = \mu^2_X\left(e^{\left[\sigma^2_{\ln(X)}\right]} - 1\right)$$

When the Lognormal distribution is defined upon an a value, this is called the three-parameters Lognormal distribution or the **Gibrat-Galton distribution**, that is, $LN(a, \mu_{\ln(X)}, \sigma^2_{\ln(X)})$. Accordingly, the transformation becomes $Y = \ln(X - a)$ so the pdf is:

$$f_X(x) = \frac{1}{(x-a)\sigma_{\ln(X-a)}\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left[\frac{\ln(X-a) - \mu_{\ln(X-a)}}{\sigma_{\ln(X-a)}}\right]^2\right]}, \text{ for } a \leq x < \infty$$

where the mean is:

$$\mu_{\ln(X-a)} = \ln(\mu_X - a) - \frac{1}{2}\ln\left[1 + \left(\frac{\sigma_X}{\mu_X - a}\right)^2\right]$$

and the variance is:

$$\sigma^2_{\ln(X-a)} = \ln\left[1 + \left(\frac{\sigma_X}{\mu_X - a}\right)^2\right]$$

Some discrete and continuous probability density functions

Distribution	Probability density/mass	Expected value	Variance	Example of Hydrological application
Binomial $B(N,p)$	$\binom{N}{n} p^n (1-p)^{N-n}$ $n = 0,1,2,..,N$	Np	$Np(1-p)$	The number n of flood events with probability p occurring in N time steps
Geometric $G(p)$	$(1-p)^{n-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	The number of time steps until a flood event with probability p occurs.
Poisson $P(\lambda)$	$\frac{e^{-\lambda}\lambda^n}{n!}$	λ	λ	The number of rain storms occurring in a given time period.
Uniform $U(a,b)$	$\frac{1}{b-a} \quad a \geq z \geq b$	$\frac{b-a}{2}$	$\frac{(b-a)^2}{12}$	(Non-informative) prior distribution of a hydrological parameter provided to a parameter estimation method
Exponential $E(\lambda)$	$\lambda e^{-\lambda z}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	The time between two rain storms
Gaussian/Normal $N(\mu,\sigma)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-[\frac{1}{2}(z-\mu)^2/\sigma^2]}$	μ	σ^2	Many applications: prior distribution for parameter optimisation; modelling of errors; likelihood functions
logNormal $L(\mu,\sigma)$	$\frac{1}{\sqrt{2\pi}\sigma z} e^{-[\frac{1}{2}(\ln z-\mu)^2/\sigma^2]}$	μ	σ^2	Hydraulic conductivity
Gamma $\Gamma(n,\lambda)$ (note: $n \in \Re$)	$\frac{\lambda^n}{\Gamma(n)} z^{n-1} e^{-\lambda z}$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	Sum of n independent random variables that are exponentially distributed with parameter λ; Instantaneous unit hydrograph of n linear reservoirs in series; pdf of travel times in a catchment;

Some discrete and continuous probability density functions

				very flexible distribution for strictly positive variables.
Beta $\beta(p,q)$	$\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} z^{p-1}(1-z)^{q-1}$ $p>0, q>0, \quad 0 \le z \le 1$	$\frac{p}{p+q}$	$\frac{pq}{(p+q)^2(p+q+1)}$	Very flexible distribution for variables with upper and lower boundaries; used as a priori distribution in Bayesian analysis and parameter estimation
Gumbel $G(a,b)$ (Extreme value distribution Type I)	$be^{-b(z-a)} \exp(-e^{-b(z-a)})$	$a+\frac{0.5772}{b}$	$\frac{\pi^2}{6b^2}$	Yearly maximum discharge used for design of dams and dikes
Weibull $W(\lambda,\beta)$ (Extreme value distribution type III)	$\lambda^{\beta} \beta z^{\beta-1} \exp[-(\lambda x)^{\beta}]$	$\frac{1}{\lambda}\Gamma(1+\frac{1}{\beta})$	$\frac{1}{\lambda^2}(A-B)$ $A=\Gamma(1+\frac{2}{\beta})$ $B=\left(\Gamma(1+\frac{1}{\beta})\right)^2$	Yearly minimum discharge used in low flow analysis.

Characteristics functions of some probability distributions

Distribution	Probability generating function	Moment generating function	Characteristic function
Binomial $B(n,p)$	$(1-p+ps)^n$	$(1-p+pe^s)^n$	$(1-p+pe^{is})^n$
Geometric $G(p)$	$\frac{ps}{1-(1-p)s}$	$\frac{pe^s}{1-(1-p)e^s}$	$\frac{pe^{is}}{1-(1-p)e^{is}}$
Poisson $P(\lambda)$	$e^{\lambda(s-1)}$	$e^{\lambda(e^s-1)}$	$e^{\lambda(e^{is}-1)}$
Uniform $U(a,b)$	-	$\frac{e^{bs}-e^{as}}{s(b-a)}$	$\frac{e^{ibs}-e^{ais}}{is(b-a)}$
Exponential $E(\lambda)$	-	$\frac{\lambda}{\lambda-s}$	$\frac{\lambda}{\lambda-is}$
Gaussian/normal $N(\mu,\sigma)$	-	$e^{\mu s+\frac{1}{2}\sigma^2 s^2}$	$e^{i\mu s-\frac{1}{2}\sigma^2 s^2}$
Gamma $\Gamma(n,\lambda)$	-	$\left(\frac{\lambda}{\lambda-s}\right)^n$	$\left(\frac{\lambda}{\lambda-is}\right)^n$

Multivariate distributions

Here, the multivariate type of distribution is examined for jointly distributed random variables. For instance the exponentially distributed storm intensity and duration lead to a bivariate exponential distribution or intensity and duration considering the correlation between these two variables. While the bivariate normal distribution is examined in detail other types are discussed briefly.

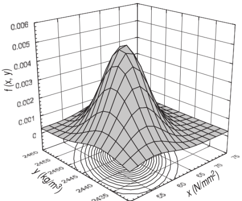
The **Bivariate normal** distribution

The joint distribution of two random variables X and Y, each normally distributed is called the **bivariate normal distribution**. The pdf in standardized form is given for Z_1 and Z_2 as:

$$f_{Z_1,Z_2}(z_1,z_2) = \left[2\pi\left(1-\rho^2\right)^{1/2}\right]^{-1} e^{\left[\frac{-\left(z_1^2-2\rho z_1 z_2+z_2^2\right)}{\left(2-2\rho^2\right)}\right]}, \text{ for } -\infty < z_1, z_2 < \infty$$

where r, $-1 \leq \rho \leq 1$, is linear correlation coefficient between the two variables. Note also that $Z_1 = \frac{X-\mu_X}{\sigma_X}$ and $Z_2 = \frac{Y-\mu_Y}{\sigma_Y}$ where $-\infty < \mu_X, \mu_Y < \infty$ and $\sigma_X, \sigma_Y > 0$. In terms of X and Y $(-\infty < x, y < \infty)$ the **bivariate normal distribution** is:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\left[-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{x-\mu_X}{\sigma_X}\frac{y-\mu_Y}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right]}$$



This figure shows the relationship of compressive strength and density in concrete. In general, the bivariate distribution shows how upon one variable is possible to predict the performance of another.

The **Bivariate exponential** distribution

There are many others useful bivariate distribution, this is case of the **bivariate exponential distribution** whose cdf is:

$$F_{X,Y}(x,y) = 1 - e^{-ax} - e^{-by} - e^{-ax-by-cxy}, \text{ for } x, y \geq 0$$

where $a, b, c > 0$ are parameters of the distribution.

The **Bivariate logistic** distribution

The cdf of the **bivariate logistic distribution** in its basic form is:

$$F_{X,Y}(x,y) = \left(1 + e^{-x} + e^{-y}\right)^{-1}, \text{ for } x, y \geq 0$$