Sums of Random Variables

Richard C. Hendriks

TU Delft
Delft University of Technology

PDF of the Sum of Two Random Variables

Consider two random variables X and Y. What is the pdf of $W = X + Y$?

TU Delft

Remember Lecture 1 for scalar RVs: Derived Random Variables

- The PDf of an arbitrary derived random variable $Y = g(X)$ is often difficult to calculate. A general procedure is to
  1. Find the CDF of $F_Y(y) = P[Y \leq y]$
  2. Compute the PDF by calculating the $f_Y(y) = \frac{dF_Y(y)}{dy}$
- Special case is a linear transformation:

$$Y = aX + b \Leftrightarrow f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

TU Delft

PDF of the Sum of Two Random Variables

$$f_W(w) = \int_{-\infty}^{\infty} f_{X,Y}(x, w-x)dx = \int_{-\infty}^{\infty} f_{X,Y}(w-y, y)dy$$

Proof: Make use of knowledge on how to derive pdf of derived RVs!

$$F_W(w) = P[X + Y \leq w] = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{w-x} f_{X,Y}(x,y)dy\right) dx$$

PDF is given by the derivative of the CDF:

$$\begin{aligned} f_W(w) &= \frac{dF_W(w)}{dw} = \int_{-\infty}^{\infty} \left(\frac{d}{dw}\left(\int_{-\infty}^{w-x} f_{X,Y}(x,y)dy\right)\right) dx \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x, w-x)dx \end{aligned}$$

Use differentiation under the integral sign:

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,t)dt = f(x,b(x))b'(x) - f(x,a(x))a'(x) + \int_{a(x)}^{b(x)} \frac{d}{dx} f(x,t)dt$$

TU Delft

PDF of the Sum of Two Independent Random Variables

For independent RVs: $f_{X,Y} = f_X f_Y$

So, for two independent RVs X and Y we get

$$\begin{aligned} f_W(w) &= \int_{-\infty}^{\infty} f_{X,Y}(x, w-x)dx \\ &= \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)dx, \end{aligned}$$

In words: the PDF of the sum of two independent RVs is the convolution of the two PDFs. (Equivalent for discrete RVs).

TU Delft

Expected value of sums of Random Variables

Consider the sum $W_n = X_1 + W_2 + \cdots + X_n$. The expected value $E[W_n]$ is given by

$$E[W_n] = E[X_1] + E[W_2] + \cdots + E[X_n]$$

The variance of W_n is given by

$$Var[W_n] = \sum_{i=1}^{n}\sum_{j=n}^{n} Cov[X_i, X_j] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov[X_i, X_j]$$

For uncorrelated variables we thus get $Var[W_n] = \sum_{i=1}^{n} Var[X_i]$

TU Delft

PDF of the Sum of n Independent Random Variables

What about the pdf of the sum of n independent variables?
For two independent RVs X and Y we have

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)dx.$$

Calculating such convolutional integrals for large n is tricky...

However, from system theory we know that convolutions in time domain correspond to multiplications in the frequency domain.

TU Delft

PDF of the Sum of n Independent Random Variables

In probability theory we can also use transforms to replace the convolution between pdfs by the multiplication of their frequency transforms.

The transform of a PDF or PMF is called: the moment generating function (if Laplace transform is used) or, the characteristic function (if Fourier transform is used).

TU Delft

Moment generating and characteristic function

The characteristic function is the (inverse) Fourier transform of the pdf:

$$\phi_X(u) = \int_{-\infty}^{\infty} f_X(x)e^{jux}dx = E[e^{jux}].$$

- $u \in \mathbb{R}$ is thus the "frequency".

Equivalently we can use the moment generating function (MGF), which is defined as the (inverse) Laplace transform of the pdf for real s:

$$\phi_X(s) = \int_{-\infty}^{\infty} f_X(x)e^{sx}dx = E[e^{sx}].$$

- In this lecture we will follow the book and use the MGF.

TUDelft

Moment generating function: Properties

For continues RVs:

$$\phi_X(s) = \int_{-\infty}^{\infty} f_X(x)e^{sx}dx = E[e^{sX}].$$

For discrete RVs

$$\phi_X(s) = \sum_{x_i \in S_x} P_X(x_i)e^{sx_i} = E[e^{sX_i}].$$

- $\phi_X(0) = E[e^0] = 1$
- The MGF is extremely useful to calculate the moments $E[X^n]$ of a RV: $E[X^n] = \frac{d^n\phi_X(s)}{ds^n}\Big|_{s=0}$.

TUDelft

Moment generating function: Properties

Proof: $E[X^n] = \frac{d^n \phi_X(s)}{ds^n}\Big|_{s=0}$.

$$\begin{aligned}
\frac{d^n \phi_X(s)}{ds^n}\Big|_{s=0} &= \frac{d^n}{ds^n} \int_{\infty}^{\infty} e^{sx} f_X(x) dx \Big|_{s=0} \\
&= \int_{\infty}^{\infty} x^n e^{sx} f_X(x) dx \Big|_{s=0} \\
&= \int_{\infty}^{\infty} x^n f_X(x) dx \\
&= E[X^n]
\end{aligned}$$

Example

Let X be exponentially distributed with MGF $\phi(s) = \frac{\lambda}{\lambda - s}$

- $E[X] = \frac{d\phi_X(s)}{ds}\Big|_{s=0} = \frac{\lambda}{(\lambda-s)^2}\Big|_{s=0} = \frac{1}{\lambda}$
- $E[X^2] = \frac{d^2\phi_X(s)}{ds^2}\Big|_{s=0} = \frac{2\lambda}{(\lambda-s)^3}\Big|_{s=0} = \frac{2}{\lambda^2}$
- $E[X^n] = \frac{d^n\phi_X(s)}{ds^n}\Big|_{s=0} = \frac{n!\lambda}{(\lambda-s)^n+1}\Big|_{s=0} = \frac{n!}{\lambda^n}$

MGF of Linearly transformed RVs

The MGF of $Y = aX + b$ is $\phi_Y(s) = e^{sb}\phi_X(as)$.

Proof:

$$\phi_Y(s) = E[e^{s(aX+b)}] = e^{sb}E[e^{saX}] = e^{sb}\phi_X(as)$$

TUDelft

The MGF for sums of RVs

Consider the sum of independent RVs $X_1, ...x_n$. The moment generating function of

$$W = \sum_{i=1}^{n} X_i$$

is then given by

$$\phi_W(s) = E[e^{sW}] = E[e^{s\sum_{i=1}^{n} X_i}] = E\left[\prod_{i=1}^{n} e^{sX_i}\right] = \prod_{i=1}^{n} \phi_{X_i}(s)$$

TUDelft

Example

Let $K_1, K_2, ...$ denote a sequence of iid Bernoulli Rvs.
Let $M = K_1 + \cdots + K_n$.

- Find MGF $\phi_K(s)$.
- Find MGF $\phi_M(s)$.
- $E[M]$ and $E[M^2]$

Example

Let $K_1, K_2, ...$ denote a sequence of iid Bernoulli Rvs.
Let $M = K_1 + \cdots + K_n$.

- Find MGF $\phi_K(s)$:
  $\phi_K(s) = E[e^{Ks}] = (1-p)e^0 + pe^s = 1 - p + pe^s$.
- Find MGF $\phi_M(s)$:
  $\phi_M(s) = \prod_{i=1}^{n} \phi_{K-i}(s) = (1 - p + pe^s)^n$
- $E[M]$ and $E[M^2]$:
  $E[M] = \frac{d}{ds} (1 - p + pe^s)^n \mid_{s=0} = npe^s (1 - p + pe^s)^{n-1} \mid_{s=0} =$
  np
  $E[M^2] = \frac{d^2}{ds^2} npe^s (1 - p + pe^s)^{n-1} \mid_{s=0} = n(n-1)(pe^s)^2 (1 - p + pe^s)^{n-2} +$
  $npe^s(1 - p + pe^s)^{n-1} \mid_{s=0} = n(n-1)p^2 + np$

The sum of Gaussian Rvs

Let $X_1, X_2, ...X_n$ denote a sequence of Gaussian Rvs.
What is the distribution of $W = X_1 + X_2 + ... + X_n$?

$$\begin{aligned} \phi_W(s) &= \phi_{X_1}(s)\phi_{X_2}(s)\dots\phi_{X_n}(s) \\ &= e^{s\mu_1+\sigma_1^2 s^2/2} e^{s\mu_2+\sigma_2^2 s^2} \dots e^{s\mu_n+\sigma_n^2 s^2} \\ &= e^{s(\mu_1+\mu_2+\cdots+\mu_n)+(\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2)s^2/2}. \end{aligned}$$

Distribution of W is thus again Gaussian with mean $\mu_1+\mu_2+\cdots+\mu_n$ and variance $\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$.

TU Delft

The Central Limit Theorem

Given a sequence of iid random variables $X_1, X_2, \dots, X_n$, each with expected value μ_X and variance σ_X^2.

The CDF of $Z_n = (\sum_{i=1}^n X_i - n\mu_X)/\sqrt{n\sigma_X^2}$ then has the property:

$$\lim_{n\to\infty} F_n = \Phi(z).$$

In other words, if n becomes "large", the distribution of the sum of iid random variables approaches a Gaussian distribution.

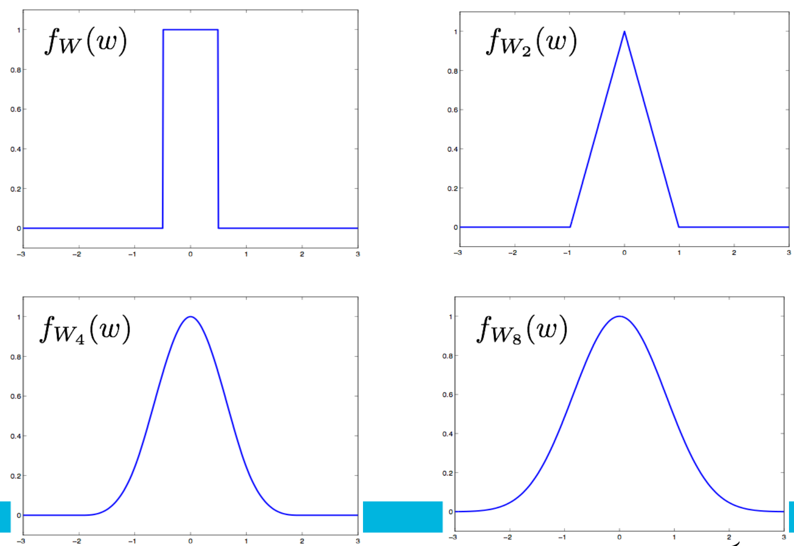
TU Delft

The Central Limit Theorem

- Although the theorem states that $\lim_{n\to\infty}$, it already holds for a relatively small number of random variables.
- There do exist "weaker" versions of the central limit theorem where the sum should consist not even of IID Rvs.

TU Delft

The Central Limit Theorem: illustration



TU Delft