ECE 2521: Analysis of Stochastic Processes

Lecture ~~9~~ 8

Department of Electrical and Computer Engineering
University of Pittsburgh

November, 3th 2021

Azime Can-Cimino

Random Vectors: $X = [x_1, \dots, x_N]$

$$F_X(x_1, \dots, x_N) = P\{X_1 \le x_1, \dots, X_N \le x_N\}$$

$$f_X(x_1, \dots, x_N) = \frac{d^n}{dx_1 \cdots dx_N} F_X(x_1, \dots, x_N)$$

$$P(X \in A) = \int f_X(x)\, dx$$

n-dim. over the region A

$$\left.\begin{aligned} G_1(x) &= y_1 \\ &\vdots \\ G_n(x) &= y_n \end{aligned}\right\} n \times n$$

$$f_{y_1, y_2, \dots, y_n}(y_1, y_2, \dots, y_n) = 0$$

(Joint PDF)

Single solution: $X = [x_1, \dots, x_N]$

$$f_Y(y_1, y_2, \dots, y_N) = \frac{f_X(x_1, \dots, x_N)}{|J(x_1, \dots, x_N)|}$$

(Jacobian)

$y = x A_{N\times N}$ A^{-1} exists

(Linear transf.)

$$\boxed{\eta_y = \eta_x \cdot A}$$

$\eta_x = E[x] = [E(x_1), E(x_2), \dots, E(x_N)]$

$$J = \det \begin{vmatrix} \frac{dy_1}{dx_1} & \dots & \frac{dy_N}{dx_1} \\ \vdots & & \\ \frac{dy_N}{dx_1} & \dots & \frac{dy_N}{dx_N} \end{vmatrix} = \det(A)$$

covariance

$$C_y = E\{[y - \eta_y]^T [y - \eta_y]\}$$

$$= E\{[(x - \eta_x)A]^T [(x - \eta_x)A]\}$$

$$C_y = A^T C_x A$$

correlation

$$R = E(x^T x) = E\left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} [x_1 \dots x_N] \right\}$$

$$= \begin{bmatrix} E(x_1^2) & E(x_1 x_2) & \dots & E(x_1 x_N) \\ & \ddots & & \\ E(x_N x_1) & & & E(x_N^2) \end{bmatrix}$$

$$C = E\{[x - \mu_x]^T [x - \mu_x]\}$$

Take inverse?

$$= \begin{bmatrix} \sigma_{x_1}^2 & \mathrm{cov}(x_1, x_2) & \dots & \mathrm{cov}(x_1, x_N) \\ & & & \\ & & & \sigma_{x_N}^2 \end{bmatrix} = R - \eta_x^T \eta_x$$

want to take C^{-1}

Note: $y = G(x)$ $\quad \mathbb{R}^n \to \mathbb{R}$

Scalar case

$$E(y) = \int_{-\infty}^{\infty} y \cdot f_y(y)\, dy = \int_{-\infty}^{\infty} \int \cdots \int_{-\infty}^{\infty} G(x) f_x(x)\, dx$$

Vector case

$$E[y] = \left[E[y_1], \ldots, E[y_N] \right]$$

Independent: PDF and CDF can be written as factors

Ex: $x_i \sim f_i(x_i) \quad i = 1, \dots, n \quad$ (independent)

$$y_k = x_1 + \dots + x_k \qquad k = 1, \dots, n$$

$$\left.\begin{aligned} x_1 &= y_1 \\ x_1 + x_2 &= y_2 \\ &\vdots \\ x_1 + \dots + x_n &= y_n \end{aligned}\right\}$$

unique solution

$$x_k = y_k - y_{k-1}$$

$$J = 1$$

$$f_y(y_1, \dots, y_n) = f_1(y_1) f_2(y_2 - y_1) \dots f_n(y_n - y_{n-1})$$

Note

$$f(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3) \quad \text{for independence}$$

$$f(x_1, x_2) = f(x_1) f(x_2)$$

$$f(x_1, x_2) = f(x_1) f(x_2)$$

$$f(x_1, x_3) = f(x_1) f(x_3)$$

$$f(x_2, x_3) = f(x_2) f(x_3)$$

$$\Downarrow$$

$$f(x_1, x_2, x_3) \neq f(x_1) f(x_2) f(x_3)$$

Linear Transformation of jointly Gaussian RVs

$$y = AX \longrightarrow x = YA^{-1}$$

$$\left.\begin{aligned} \eta_y &= \eta_x \cdot A \\ C_y &= A^T C_x A \end{aligned}\right\} \quad f_y(y) = \frac{f_x(YA^{-1})}{\det(A)}$$

$$= \frac{1}{(2\pi)^{n/2}\sqrt{\det C_x}} \frac{1}{\det(A)} \exp\left\{-\frac{1}{2}(yA^{-1} - \eta_x)(yA^{-1} - \eta_x)^T\right\}$$

$$= \frac{1}{(2\pi)^{n/2}\sqrt{\det C_x} \det|A|} \exp\left\{\frac{-1}{2}(y - \eta_y)\underline{A^{-1}C_x^{-1}(A^{-1})^T}(y - \eta_y)^T\right\}$$

$$A^{-1}C_x^{-1}(A^{-1})^T = (A^T C_x A)^{-1} = C_y^{-1}$$

$W = CZ + \alpha$ $\qquad$ $Z \sim N(\eta_z, K_z)$

$\Rightarrow W \sim N(C\eta_z + \alpha,\ CK_zC^T)$

<u>Proof:</u> $\phi_w(\omega) = E\left[e^{J\omega^T w}\right] = E\left[e^{J\omega^T(Cz+\alpha)}\right]$

$= e^{J\omega^T\alpha}\, E\left[e^{JC^T\omega^T z}\right]$

$= e^{J\omega^T\alpha}\, \Phi_z(C^T\omega)$

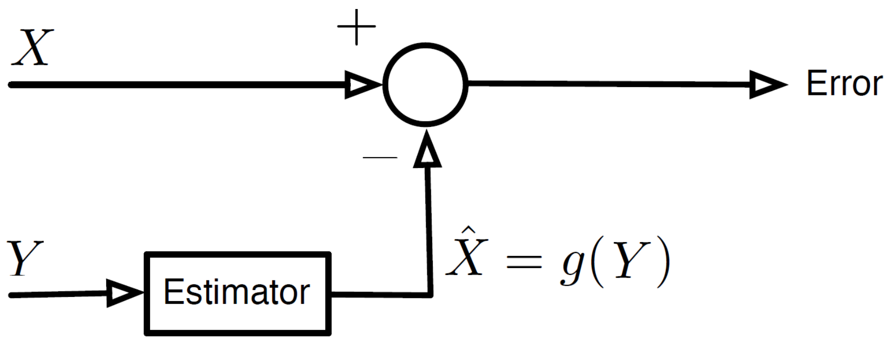
$\Phi_z(C^T\omega) = \exp\left\{J\cdot(\omega^T C\eta_z) - \frac{1}{2}(\omega^T C)K_z(C^T\omega)\right\}$

$\phi_w(\omega) \Rightarrow N(\alpha + C\eta_z,\ CK_zC^T)$

Estimation of Random Variables

- Estimating the parameters of one or more random variables (e.g. probabilities, means, variances, or covariances)
- Estimating the value of an inaccessible random variable X in terms of the observation of an accessible random variable Y:
  - Prediction Problems: predict future based on current and past observations
  - Interpolation Problems: given samples of a signal, we wish to interpolate to some in-between point in time
  - Filtering Problems: filter the noise out of a sequence of observations to provide the best estimate of the desired signal

Mean-Square Estimation (MSE)



- Assume X and Y are correlated to some degree
- If Y is observed, then estimate X so as to minimize the mean-square error:

$$e = E[(X - g(Y))^2]$$

Constant MSE

(a) Estimate the random variable X by a constant $\hat{X} = g(Y) = a$ so that the mean-square error is minimized.

(b) What is the mean-square error for this estimator?

Linear MSE

- Estimate X by a linear function $g(Y) = aY + b$ so that the mean-square error is minimized:

$$\min_{a,b} E[(X - aY - b)^2]$$

Step 1 We can apply the result from the previous example if we view the problem as estimating the random variable $(X - aY)$ with a constant b, such that:

$$b^* = E[X - aY] = E[X] - aE[Y]$$

Step 2 The minimization problem simplifies to one parameter a:

$$\min_{a} E[(X - E[X] - a(Y - E[Y]))^2]$$

such that $a^* = \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}$

Linear MSE

Normalization

- The linear estimate $g(Y) = \overbrace{aY + b}^{\text{nonhomogeneous}}$ of X is obtained:

$$\boxed{\hat{X} = E[X] + \mathrm{Cov}(X, Y)\frac{Y - E[Y]}{\mathrm{Var}(Y)}}$$

Note The linear mean-square estimator depends on second order moments: mean, variance and covariance.

- The minimum error of the linear MSE:
$\epsilon_{MIN} = \mathrm{Var}(X)\left(1 - \rho^2\right)$.

Correlation coefficient

Linear MSE

- Knowing the correlation coefficient $\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$, the linear estimate $g(Y) = aY + b$ of X can be rewritten as:

$$\hat{X} = E[X] + \rho\sqrt{\mathrm{Var}(X)}\frac{Y - E[Y]}{\sqrt{\mathrm{Var}(Y)}}$$

- $E[X]$ provides the mean value of the random variable being estimated
- The term $\frac{Y-E[Y]}{\sqrt{\mathrm{Var}(Y)}}$ is a zero-mean, unit-variance version of Y
- Multplying this term by $\sqrt{\mathrm{Var}(X)}$ rescales Y to yield the variance of the random variable being estimated
- The correlation coefficient ρ specifies the sign and extent of the estimate

Quiz What if X and Y are not correlated?

Orthogonality of the Linear MSE

- Recall that the minimization of the mean-square error to obtain a^* yields:

$$E[(X - E[X] - a^*(Y - E[Y]))\,(Y - E[Y])] = 0$$

  where the optimal linear MSE is given by $\hat{X} = E[X] - a^*(Y - E[Y])$.
- The **orthogonality principle** states that the error of the best linear estimator is orthogonal to the observation $Y - E[Y]$.

Mean-Square Estimation (MSE)

Nonlinear estimator

- Estimator: $\hat{X} = g(Y)$
- Find $g(.)$ such that it minimizes $E[(X - g(Y))^2]$
- Solution: $\hat{X} = E[X|Y]$

Remarks

- $E[X|Y]$ is in general a nonlinear function of Y (nonlinear estimator)
- If X and Y are independent, then $E[X|Y] = E[X]$
- The minimum error $\epsilon_{MIN} = E[(X - E[X|Y])^2]$ is the conditional variance of X given Y
- $g^*(Y) = E[X|Y]$ is the best approximation in the mean-square sense of X among all possible functions, or $E[(X - g(Y))^2] \geq E[(X - E[X|Y])^2]$ for all functions $g(.)$
- If X and Y are Gaussian, then $E[X|Y]$ is a linear function of Y

Estimation using a Vector of Observations

- Estimator: $\hat{X} = g(\mathrm{Y})$ where $\mathrm{Y} = [Y_1, Y_2, ..., Y_n]^T$ is a vector
- Find $g(.)$ such that it minimizes $E[(X - g(\mathrm{Y}))^2]$
- Solution: $\hat{X} = E[X|\mathrm{Y}]$
- Linear MSE:

(i) $\hat{X} = g(\mathrm{Y}) = \mathrm{a}^T\mathrm{Y} = \sum_{k=1}^{n} a_k Y_k$ and $E[X] = E[Y] = 0$

  - $E[X\mathrm{Y}] = \mathrm{R}_Y\mathrm{a}$ such that $\mathrm{a} = \mathrm{R}_Y^{-1}E[X\mathrm{Y}]$, where R_Y is the correlation matrix
  - $\epsilon_{MIN} = E[X^2] - \mathrm{a}^T E[\mathrm{Y}X] = \mathrm{Var}[X] - \mathrm{a}^T E[\mathrm{Y}X]$

(ii) $\hat{X} = \mathrm{a}^T\mathrm{Y} + b = \sum_{k=1}^{n} a_k Y_k + b$ and $E[X] = \mu_X$, $E[Y] = \mu_Y$

  - $b^* = E[X] - \mathrm{a}^T\mu_Y$
  - Therefore $\hat{X} = \mathrm{a}^T(\mathrm{Y} - \mu_Y) + \mu_X$ such that: $\hat{X} - \mu_X = W = \mathrm{a}^T\mathrm{Z}$
  - $\mathrm{a}^* = \mathrm{R}_Z^{-1}E[W\mathrm{Z}] = \mathrm{K}_Y^{-1}E[(X - \mu_X)(\mathrm{Y} - \mu_Y)]$, where K_Y is the covariance matrix
  - $\epsilon_{MIN} = \mathrm{Var}[X] - \mathrm{a}^T E[(X - \mu_X)(\mathrm{Y} - \mu_Y)]$

Sums of Random Variables

$Z = X + Y$

- In Chapter 7, we will study the properties of the sums of random variables such as the mean, variance, and the PDF of the sum
- In deriving the PDF of the sum of random variables, we will use tools such as the Moment Generating Functions
- Let $X_1, X_2, \cdots, X_n$ be random variables and W_n be their sum:

$$W_n = X_1 + X_2 + \cdots + X_n$$

Expected Value of the Sum of Random Variables

$$E[W_n] = E[X_1] + E[X_2] + \cdots + E[X_n] = \sum_{i=1}^{n} E[X_i]$$

- The expected value of sum is equal to the sum of individual expected values.

Variance of the Sum of Random Variables

- Let us first look at the simple case $n = 2$: For the sum of two random variables $W_2 = X_1 + X_2$:

$$\begin{aligned}
\text{Var}[W_2] &= \overbrace{E\left[(W_2 - E[W_2])^2\right]}^{\text{Variance def.}} = E\left[(X_1 + X_2 - E[X_1 + X_2])^2\right] \\
&= E\left[(X_1 + X_2 - E[X_1] - E[X_2])^2\right] \\
&= E\left[(X_1 - E[X_1])^2 + (X_2 - E[X_2])^2 + 2(X_1 - E[X_1])(X_2 - E[X_2])\right] \\
&= E\left[(X_1 - E[X_1])^2\right] + E\left[(X_2 - E[X_2])^2\right] + 2E[(X_1 - E[X_1])(X_2 - E[X_2])] \\
&= \text{Var}[X_1] + \text{Var}[X_2] + 2\text{Cov}[X_1, X_2]
\end{aligned}$$

Variance of the Sum of Random Variables

- The general case $W_n = X_1 + X_2 + \cdots + X_n$:

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

- In general, the variance of sum is not equal to the sum of individual variances (since we also need to know the co-variances)
- **Special case**: When $X_1, \cdots, X_n$ are *uncorrelated* then:

(independent)

$$\text{Var}[W_n] = \sum_{i=1}^{n} \text{Var}[X_i]$$

- Recall that two random variables X_i and X_j are **uncorrelated** if $\text{Cov}[X_i, X_j] = 0$.

Example 1

- Let $X_1, X_2, \cdots, X_n$ be **independent** and **identically distributed** (i.i.d) random variables, each with mean μ and variance σ^2. Find the expected value and the variance of $W_n = X_1 + X_2 + \cdots + X_n$.

Solution: The mean is computed as follows:

$$E[W_n] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \mu = n\mu$$

- Since any two independent random variables are uncorrelated, their covariance is equal to zero:

Add variances

$$\mathrm{Var}[W_n] = \sum_{i=1}^{n} \mathrm{Var}[X_i] = \sum_{i=1}^{n} \sigma^2 = n\sigma^2$$

Example 2

- Let $X_1, \cdots, X_n$ be random variables, each with mean μ and covariance function:

$$\mathrm{Cov}\left[X_i, X_j\right] = \sigma^2 \rho^{-|i-j|},$$

where $|\rho| < 1$. Find the mean and the variance of $Y_i = X_i + X_{i+1} + X_{i+2}$.

$\mu_y = 3\mu$

$cov(x_i, x_{i+1}) = \sigma^2 \rho^{-1}$

$var_y = 3\sigma^2 + 2\sigma^2 \left(\frac{2}{\rho} + \frac{1}{\rho^2}\right)$

$cov(x_{i+1}, x_{i+2}) = \sigma^2 \rho^{-1}$

$= \sigma^2 \left(3 + \frac{4}{\rho} + \frac{2}{\rho^2}\right)$

$cov(x_i, x_{i+2}) = \sigma^2 \rho^{-2}$

PDF of Sum of Two Random Variables

- The PDF of the sum of two random variables $W = X + Y$ is:

convolution

$$f_W(w) = \int_{-\infty}^{\infty} f_{X,Y}(x, w-x)dx = \int_{-\infty}^{\infty} f_{X,Y}(w-y, y)dy \qquad (1)$$

- **Special Case:** When X and Y are **independent** random variables (i.e. $f_{X,Y}(x,y) = f_X(x)f_Y(y)$), the PDF of $W = X + Y$ is:

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

Recall: The convolution of two functions $f(t)$ and $g(t)$:

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$

- When X and Y are **independent** random variables, the PDF of $W = X + Y$ is the **convolution** of the marginal PDFs $f_X(x)$ and $f_Y(y)$: $f_W(w) = f_X(x) * f_Y(y)$.

Review: Graphical calculation of convolutions

- The graphical evaluation of the convolution:

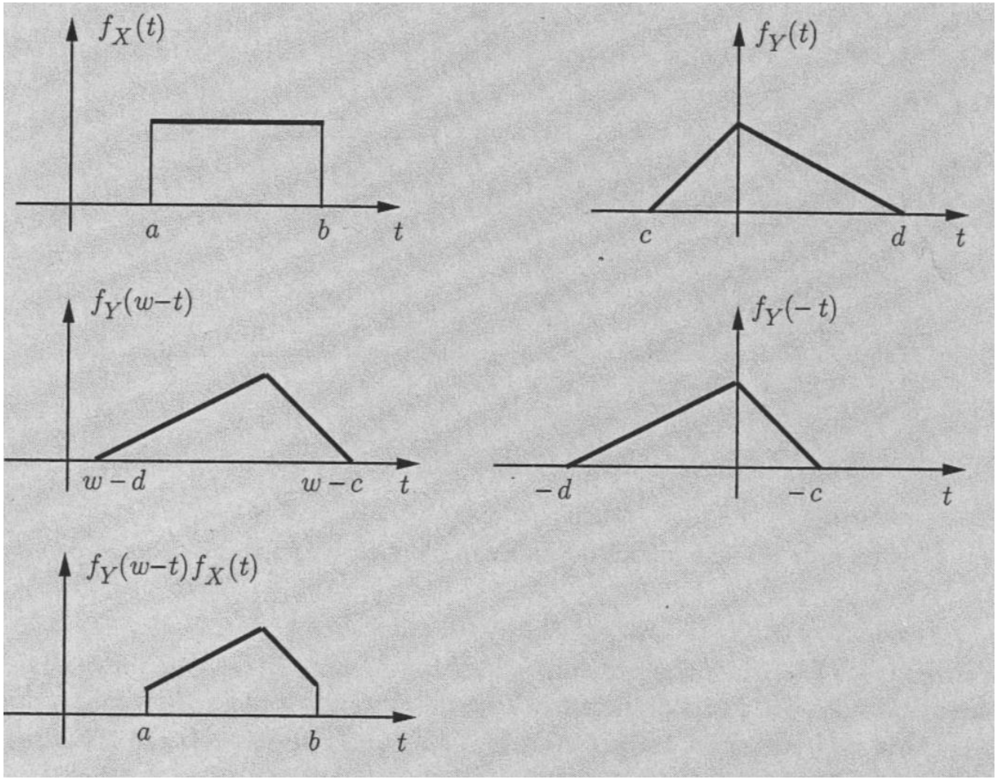
$$f_W(w) = \int_{-\infty}^{\infty} f_X(t) f_Y(w-t) dt$$

consists of the following steps:

1. Plot $f_Y(w-t)$ as a function of t. This plot has the same shape as $f_Y(t)$ except that it is first "flipped" $(f_Y(-t))$ and then shifted by an amount w (i.e. $f_Y(w-t)$). If $w > 0$, this is a shift to the right, if $w < 0$ this is a shift to the left.
2. Place the plots $f_X(t)$ and $f_Y(w-t)$ on top of each other, and form their product.
3. Calculate the value of $f_W(w)$ by calculating the integral of the product of these two plots.

- By varying the amount w by which we are shifting, we obtain $f_W(w)$ for any w.

Review: Graphical calculation of convolutions



Example 1

- Let X and Y be independent random variables that are uniformly distributed in the interval $[0, 1]$. Find the PDF of $W = X + Y$.

Solution: Since X and Y are independent, the PDF of W is defined as:

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

where, X and Y are uniformly distributed, i.e.,

$$f_X(x) = f_Y(y) = \begin{cases} 1 & 0 \leq x \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

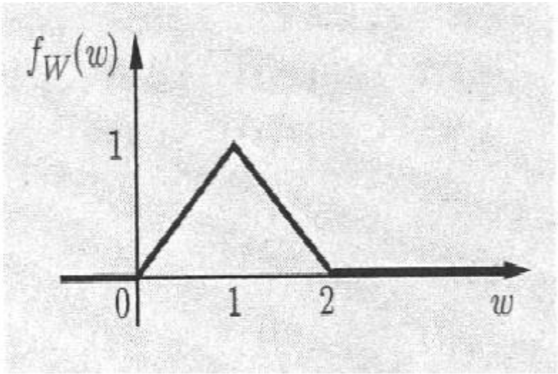
- We note that $f_X(x)$ is non-zero (and equal to 1) for $0 \leq x \leq 1$ and $f_Y(w-x)$ is also non-zero (and equal to one) for $0 \leq w - x \leq 1$ (or equivalently $w - 1 \leq x \leq w$).

Example 1 - Solution (continued)

- Combining these two inequalities, the integrand of the PDF of W (i.e., $f_X(x)f_Y(w-x)$) is non-zero for:

$$\max\{0, w-1\} \leq x \leq \min\{1, w\}$$

$$f_W(w) = \begin{cases} \min\{1, w\} - \max\{0, w-1\} & 0 \leq w \leq 2, \\ 0 & \text{otherwise} \end{cases}$$



Example 2

- Find the PDF of the sum of two zero-mean, unit-variance Gaussian random variables with correlation coefficient $\rho = -1/2$.

Solution: Let $W = X + Y$ denote the sum of the two Gaussian random variables X and Y with joint PDF:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \qquad -\infty < x, y < \infty \tag{2}$$

- Replace Eq. (2) into Eq. (1) to obtain the PDF of W:

$$\begin{aligned} f_W(w) &= \int_{-\infty}^{\infty} f_{X,Y}(x, w-x)dx \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho x(w-x)+(w-x)^2)/2(1-\rho^2)} dx \\ &= \frac{1}{2\pi\sqrt{3/4}} \int_{-\infty}^{\infty} e^{-(x^2-xw+w^2)/2(3/4)} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} \end{aligned}$$

Note: The sum of two **non-independent** Gaussian RVs is also Gaussian!

MGF for Sums of Independent Random Variables

- MGFs or transforms are useful in finding the distributions of sums of independent random variables.
- Let $X_1, X_2, \ldots, X_n$ be n **independent** random variables and let W denote their sum:

$$W = X_1 + X_2 + \cdots + X_n$$

- The MGF of W_n is given by:

$$\Phi_W(s) = \Phi_{X_1}(s)\Phi_{X_2}(s)\ldots\Phi_{X_n}(s) \tag{3}$$

$$\boxed{\text{Adding independent RVs} \Longleftrightarrow \text{Multiplication of MGFs}}$$

Proof:

$$\begin{aligned}\Phi_W(s) &= E[e^{sW}] = E[e^{s(X_1+X_2+\cdots+X_n)}] = E[e^{sX_1}e^{sX_2}\ldots e^{sX_n}] \\ &= E[e^{sX_1}]E[e^{sX_2}]\ldots E[e^{sX_n}] = \Phi_{X_1}(s)\Phi_{X_2}(s)\ldots\Phi_{X_n}(s)\end{aligned}$$

MGF for Sums of Independent Random Variables

$$\boxed{\text{Adding independent RVs} \Longleftrightarrow \text{Multiplication of MGFs}}$$

Special case: When $X_1, X_2, \ldots, X_n$ are i.i.d (**independent** and **identically distributed**), each with MGF $\Phi_{X_i}(s) = \Phi_X(s)$, then

$$\begin{aligned}\Phi_W(s) &= \Phi_{X_1}(s)\Phi_{X_2}(s)\ldots\Phi_{X_n}(s)\\ &= (\Phi_X(s))^n\end{aligned}$$

MGFs for Common Random Variables

Random Variable	PMF	MGF $\phi_X(s)$
Bernoulli (p)	$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \\ 0 & \text{otherwise} \end{cases}$	$1-p+pe^s$
Binomial (n, p)	$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$	$(1-p+pe^s)^n$
Geometric (p)	$P_X(x) = \begin{cases} p(1-p)^{x-1} & x=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{pe^s}{1-(1-p)e^s}$
Pascal (k, p)	$P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$	$(\dfrac{pe^s}{1-(1-p)e^s})^k$
Poisson (α)	$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x=0,1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$	$e^{\alpha(e^s-1)}$
Disc. Uniform (k, l)	$P_X(x) = \begin{cases} \frac{1}{l-k+1} & x=k,k+1,\ldots,l \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{e^{sk}-e^{s(l+1)}}{1-e^s}$

MGFs for Common Random Variables

Random Variable	PDF	MGF $\phi_X(s)$
Constant (a)	$f_X(x) = \delta(x-a)$	e^{sa}
Uniform (a, b)	$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{e^{bs} - e^{as}}{s(b-a)}$
Exponential (λ)	$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{\lambda}{\lambda - s}$
Erlang (n, λ)	$f_X(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$(\dfrac{\lambda}{\lambda - s})^n$
Gaussian (μ, σ)	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$	$e^{s\mu + s^2\sigma^2/2}$

The sum of n independent **Poisson** random variables is a **Poisson** random variable

- Let $X_1, \ldots, X_n$ denote n independent Poisson random variables each with $E[X_i] = \alpha_i$.
- The MGF table gives the MGF $\Phi_{X_i}(s) = e^{\alpha_i(e^s-1)}$
- Since X_is are independent, using Eq. (3):

$$\begin{aligned} \Phi_W(s) &= \Phi_{X_1}(s) \ldots \Phi_{X_n}(s) = e^{\alpha_1(e^s-1)} e^{\alpha_2(e^s-1)} \ldots e^{\alpha_n(e^s-1)} \\ &= e^{(\alpha_1+\cdots+\alpha_n)(e^s-1)} = e^{\alpha_T(e^s-1)} \end{aligned}$$

  where $\alpha_T = \alpha_1 + \cdots + \alpha_n$.
- Now using the MGF table, $\Phi_W(s)$ is the MGF of a Poisson RV
- Therefore W is also a Poisson random variable with $E[W] = \alpha_T$:

$$P_W(w) = \begin{cases} \frac{{\alpha_T}^w}{w!} e^{-\alpha_T} & w = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases}$$

The sum of n independent **Gaussian** random variables is a **Gaussian** random variable

central limit theorem?

- Let $X_1, \ldots, X_n$ denote n independent Gaussian random variables each with mean μ_i and variance σ_i^2.
- The MGF table gives the MGF $\Phi_{X_i}(s) = e^{s\mu_i + \sigma_i^2 s^2/2}$. Since X_is are independent, using Eq. (3):

$$\begin{aligned}\Phi_W(s) &= \Phi_{X_1}(s)\ldots\Phi_{X_n}(s) = e^{s\mu_1+\sigma_1^2 s^2/2}\ldots e^{s\mu_n+\sigma_n^2 s^2/2} \\ &= e^{s(\mu_1+\cdots+\mu_n)+\left(\sigma_1^2+\cdots+\sigma_n^2\right)s^2/2}\end{aligned}$$

- Now using the MGF table, $\Phi_W(s)$ is the MGF of a Gaussian random variable, with mean $\mu_1 + \cdots + \mu_n$ and variance $\sigma_1^2 + \cdots + \sigma_n^2$.

$$f_W(w) = \frac{1}{(\sigma_1^2 + \cdots + \sigma_n^2)\sqrt{2\pi}} e^{-(w-(\mu_1+\cdots+\mu_n))^2/2\left(\sigma_1^2+\cdots+\sigma_n^2\right)}$$

Example 1

- Find the PDF of a sum of n independent exponentially distributed random variables all with parameter λ.

Solution: Let $X_1, \ldots, X_n$ denote n i.i.d exponential random variables with parameter λ.

- The MGF table gives the MGF $\Phi_{X_i}(s) = \frac{\lambda}{\lambda - s}$.
- Let $W = X_1 + \cdots + X_n$ then:

$$\Phi_W(s) = (\Phi_{X_i}(s))^n = \left(\frac{\lambda}{\lambda - s}\right)^n$$

- The MGF table shows that W has the MGF of an Erlang(n, λ) random variable, i.e., W has an Erlang(n, λ) PDF.

Example 2

- Find the MGF and the PDF for a sum of n independent identically geometrically distributed random variables.

Solution: Let $X_1, \ldots, X_n$ denote n i.i.d geometric (p) random variables.

- The MGF table gives the MGF of a geometric (p) RV as:

$$\Phi_{X_i}(s) = \frac{pe^s}{1-(1-p)e^s}$$

- Let $W = X_1 + \cdots + X_n$ then:

$$\Phi_W(s) = (\Phi_{X_i}(s))^n = \left(\frac{pe^s}{1-(1-p)e^s}\right)^n$$

- The MGF table shows that W has the MGF of a Pascal(n,p) random variable, i.e., W has a Pascal(n,p) PDF:

$$P_W(w) = \binom{w-1}{n-1} p^n (1-p)^{w-n}$$

Example 3

- Find the mean and variance of a binomial random variable $W \sim \text{binomial}(n, p)$ using its MGF.

Solution: A binomial random variable $W \sim \text{binomial}(n, p)$ is the sum of n independent Bernoulli random variables X_i all with a common parameter p, i.e, $W = X_1 + \cdots + X_n$.

- The MGF of a Bernoulli (p) random variable X_i is given by:

$$\Phi_{X_i}(s) = e^{1s}p + e^{0s}(1-p) = 1 - p + pe^s$$

- Now, the MGF of $W = X_1 + \cdots + X_n$ is given by:

$$\Phi_W(s) = (\Phi_{X_i}(s))^n = (1 - p + pe^s)^n$$

Example 3 - Solution

- The mean of W is:

$$E[W] = \frac{d}{ds}\Phi_W(s)\Big|_{s=0} = \frac{d}{ds}(1-p+pe^s)^n\Big|_{s=0}$$
$$= npe^s(1-p+pe^s)^{n-1}\Big|_{s=0} = np$$

- The second moment of W is:

$$E\left[W^2\right] = \frac{d^2}{ds^2}\Phi_W(s)\Big|_{s=0} = \frac{d^2}{ds^2}(1-p+pe^s)^n\Big|_{s=0}$$
$$= npe^s(1-p+pe^s)^{n-1} + n(n-1)p^2e^{2s}(1-p+pe^s)^{n-}$$
$$= np + n(n-1)p^2$$

- The variance of W is:

$$\text{Var}[W] = E\left[W^2\right] - (E[W])^2 = np + n(n-1)p^2 - n^2p^2 = np(1-p)$$

Random Sums of Independent Random Variables

- So far we have assumed that the number of variables in the sum is known and fixed.
- Now we will consider the case where the number of random variables being added is also a random variable itself.
- In this section we consider sums of i.i.d random variables where the number of terms in the sum is also random.
- Let N be a random variable and let $X_1, X_2, \cdots, X_N$ be i.i.d random variables and assume N is independent of the X_is
- The *random sum* of random variables is:

$$R = X_1 + X_2 + \cdots + X_N$$

Example 1

- At a bus terminal, count the number of people arriving on buses during one minute, if:
  - The number of buses arriving in one minute is N (N is a random variable)
  - The number of people on the ith bus is K_i (K_is are i.i.d random variables)
- The number of people arriving in one minute is a random sum:

$$R = K_1 + K_2 + \cdots + K_N$$

Example 2

- Count the number of data packets received successfully over a communication link in one minute, if:
  - The number of data packets arriving in one minute is N (N is a random variable)
  - Each packet is either successfully decoded or not
  - Let $X_i = 0$ if packet i is not decoded and $X_i = 1$ if packet i is decoded successfully (X_is are i.i.d random variables)
- The number of data packets received successfully in one minute is a random sum:

$$R = X_1 + X_2 + \cdots + X_N$$

Example 3

- Find the execution time of all computer jobs submitted in an hour, if:
  - The number of computer jobs submitted in one hour is N (N is a random variable)
  - The execution time for job i is T_i (T_is are i.i.d random variables)
- The execution time of all computer jobs submitted in an hour is a random sum:

$$R = T_1 + T_2 + \cdots + T_N$$

Theorem

- Let:

$$R = X_1 + X_2 + \cdots + X_N$$

  where

  - N: nonnegative integer-valued random variable with MGF $\Phi_N(s)$
  - X_i: i.i.d random variables each with MGF $\Phi_X(s)$
  - N is independent of X_i's

- The MGF, mean and variance of R are:

$$\begin{aligned} \Phi_R(s) &= \Phi_N\left(\ln \Phi_X(s)\right) \\ E[R] &= E[N]\,E[X] \\ \mathrm{Var}[R] &= E[N]\,\mathrm{Var}[X] + \mathrm{Var}[N]\,(E[X])^2 \end{aligned}$$

Example 1

- Let $X_1, X_2, \ldots$ denote a sequence of i.i.d random variables with exponential PDF ($\lambda = 1$), and N denote a geometric random variable ($p = 1/5$). Let $R = X_1 + \cdots + X_N$.
  (1) Find the MGF of R.
  (2) Find the PDF of R.

Example 1 - Solution

(1) R is a random sum, i.e., is the sum of a random number of random variables:

- X_is are i.i.d exponential random variables ($\lambda = 1$): $\Phi_X(s) = \frac{1}{1-s}$
- N is a geometric random variable ($p = 1/5$): $\Phi_N(s) = \frac{\frac{1}{5}e^s}{1-\frac{4}{5}e^s}$

$$\Phi_R(s) = \Phi_N\left(\ln \Phi_X(s)\right) = \frac{\frac{1}{5}e^{\ln \Phi_X(s)}}{1 - \frac{4}{5}e^{\ln \Phi_X(s)}} = \frac{\frac{1}{5}\Phi_X(s)}{1 - \frac{4}{5}\Phi_X(s)}$$

Substituting for $\Phi_X(s)$ yields $\Phi_R(s) = \frac{\frac{1}{5}}{\frac{1}{5}-s}$

(2) From the MGF table, we note that R has the MGF of an exponential random variable ($\lambda = 1/5$):

$$f_R(r) = \begin{cases} \frac{1}{5}e^{-\frac{r}{5}} & r \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Example 2

- Jane visits a number of bookstores looking for a particular book. Any given bookstore carries the book with probability p, independent of other bookstores. At each store Jane spends a random amount of time, distributed according to an exponential (λ). She keeps visiting bookstores until she finds the book she is looking for.
- Find the mean, the variance and the PDF of the total time she spends looking for the book.

Example 2 - Solution

- Let T_i denote the time she spends at each bookstore, where T_is are independent exponential (λ) random variables. The total number of stores visited N is a geometric (p) random variable.
- Let R denote the total time, $R = T_1 + \ldots + T_N$. Since the number of stores that she visits N is a random variable and the T_i's are i.i.d random variables, R denotes a random sum.
- Using the formulas for the mean of geometric and exponential random variables:

$$E[R] = E[N]\, E[X] = \frac{1}{p} \cdot \frac{1}{\lambda}$$

- Using the formulas for the variance of geometric and exponential random variables:

$$\begin{aligned} \mathrm{Var}[R] &= E[N]\, \mathrm{Var}[T] + \mathrm{Var}[N]\, (E[T])^2 \\ &= \frac{1}{p} \cdot \frac{1}{\lambda^2} + \frac{1-p}{p^2} \cdot \frac{1}{\lambda^2} = \frac{1}{\lambda^2 p^2} \end{aligned}$$

Example 2 - Solution (continued)

- The moment generating function for a geometric (p) random variable is found as $\Phi_N(s) = \frac{pe^s}{1-(1-p)e^s}$.
- The moment generating function for an exponential (λ) random variable is found as $\Phi_X(s) = \frac{\lambda}{\lambda - s}$.
- The moment generating function of the random sum R is given by:

$$\Phi_R(s) \quad = \quad \Phi_N\left(\ln \Phi_X(s)\right) = \frac{pe^{\ln \Phi_X(s)}}{1-(1-p)e^{\ln \Phi_X(s)}} = \frac{p\Phi_X(s)}{1-(1-p)\Phi_X(s)}$$

- Replacing for $\Phi_X(s)$ we have:

$$\Phi_R(s) \quad = \quad \frac{p\frac{\lambda}{\lambda - s}}{1-(1-p)\frac{\lambda}{\lambda - s}} = \frac{p\lambda}{p\lambda - s}$$

- We recognize that this is the MGF associated with an exponential $(p\lambda)$ random variable, therefore R is an exponential $(p\lambda)$:

$$f_R(r) = \begin{cases} p\lambda e^{-p\lambda r} & r \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Note that this result indicates that the sum of a geometric number of independent exponential random variables is exponential.