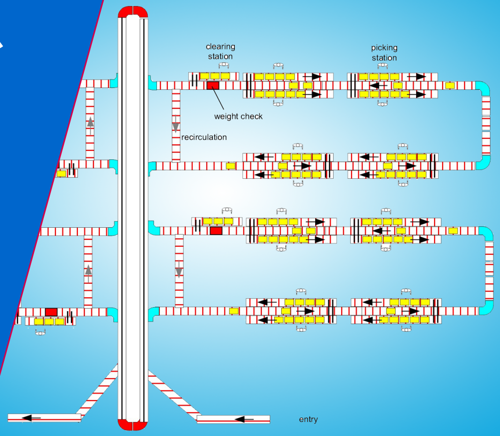
Analysis of Manufacturing Systems 4AB00

Ivo Adan



Technische Universiteit
Eindhoven
University of Technology

Contents

- Chapter 5: 5.4, 5.5
- Chapter 11: 11.1, 11.2, 11.3, 11.4.1, 11.5

Recap

- Normal random variable X with parameters μ and $\sigma > 0$,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}, \quad -\infty < x < \infty$$

  Then

$$E(X) = \mu, \quad \text{var}(X) = \sigma^2.$$

  Density $f(x)$ is denoted as $N(\mu, \sigma^2)$ density.

- Standard normal random variable X has $N(0, 1)$ density, so

$$f(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

  and

$$P(X \leq x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy.$$

Recap

- Linearity: If X is normal, then $aX + b$ is normal
- Additivity: If X and Y are independent and normal, then $X + Y$ is normal
- Probability that X lies $\geq z$ standard deviations above its mean is

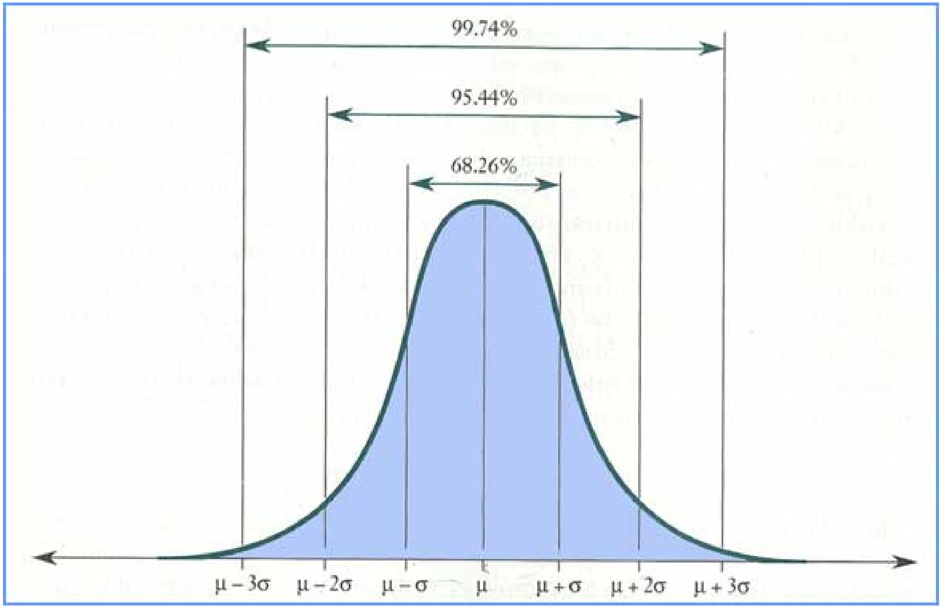
$$P(X \geq \mu + z\sigma) = 1 - \Phi(z)$$

- $100p\%$ percentile z_p of standard normal distribution is solution of

$$\Phi(z_p) = p$$

For example, $z_{0.95} = 1.64$, $z_{0.975} = 1.96$

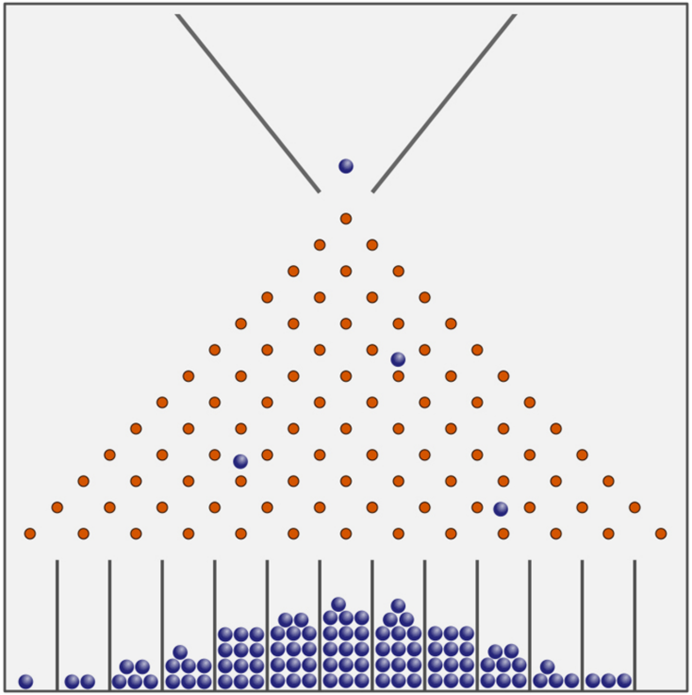
Properties of normals



Normal (or Gaussian) distribution

Central limit theorem in action

The Galton board:



Central limit theorem

$X_1, X_2, \ldots$ are independent random variables with the same distribution, and

$$\mu = E(X), \quad \sigma = \sigma(X)$$

Then

$$E(X_1 + \cdots + X_n) = n\mu, \quad \sigma(X_1 + \cdots + X_n) = \sigma\sqrt{n}$$

Question:

What is the distribution of $X_1 + \cdots + X_n$ when n is large?

Central limit theorem

For any $a < b$,

$$\lim_{n\to\infty} P\left(a \leq \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq b\right) = \Phi(b) - \Phi(a).$$

In words:

$X_1 + \cdots + X_n$ has approximately a normal distribution when n is large,
no matter what form the distribution of X_i takes!

Remark:

- Central limit theorem still valid for nonidentical X_i
- Many random quantities are addition of many small random effects:
  That is why the normal distribution often appears!

Central limit theorem

Example:

A friend claims to have tossed 5.227 heads in 10.000 tosses?

Do you believe this guy?

Confidence intervals

Question: How to estimate the unknown $\mu = E(X)$ of a random variable X?

Suppose n independent repetitions of experiment are performed, where X_k is the outcome of experiment k, $k = 1, \ldots, n$

An estimator for the unknown $\mu = E(X)$ is the sample mean

$$\overline{X}(n) = \frac{1}{n} \sum_{k=1}^{n} X_k$$

The Central limit theorem tells us

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

has an approximately standard normal distribution with $\sigma = \sigma(X)$

Confidence intervals

So

$$\frac{\overline{X}(n) - \mu}{\sigma/\sqrt{n}}$$

has an approximately standard normal distribution!

- Percentile $z_{1-\frac{1}{2}\alpha}$ is the point for which the area under the standard normal curve between points $-z_{1-\frac{1}{2}\alpha}$ and $z_{1-\frac{1}{2}\alpha}$ equals $100(1-\alpha)\%$.
- $z_{1-\frac{1}{2}\alpha}$ is 1.960 and 2.324 for $\alpha = 0.05$ and $\alpha = 0.01$.

Then

$$P\left(-z_{1-\frac{1}{2}\alpha} \leq \frac{\overline{X}(n) - \mu}{\sigma/\sqrt{n}} \leq z_{1-\frac{1}{2}\alpha}\right) \approx 1 - \alpha$$

or...

Confidence intervals

this leads to the following interval containing μ with probability $1-\alpha$

$$P\left(\overline{X}(n) - z_{1-\frac{1}{2}\alpha}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}(n) + z_{1-\frac{1}{2}\alpha}\frac{\sigma}{\sqrt{n}}\right) \approx 1-\alpha.$$

Remarks:

- If σ is unknown, it can be estimated by square root of sample variance

$$S^2(n) = \frac{1}{n}\sum_{k=1}^{n}\left[X_k - \overline{X}(n)\right]^2$$

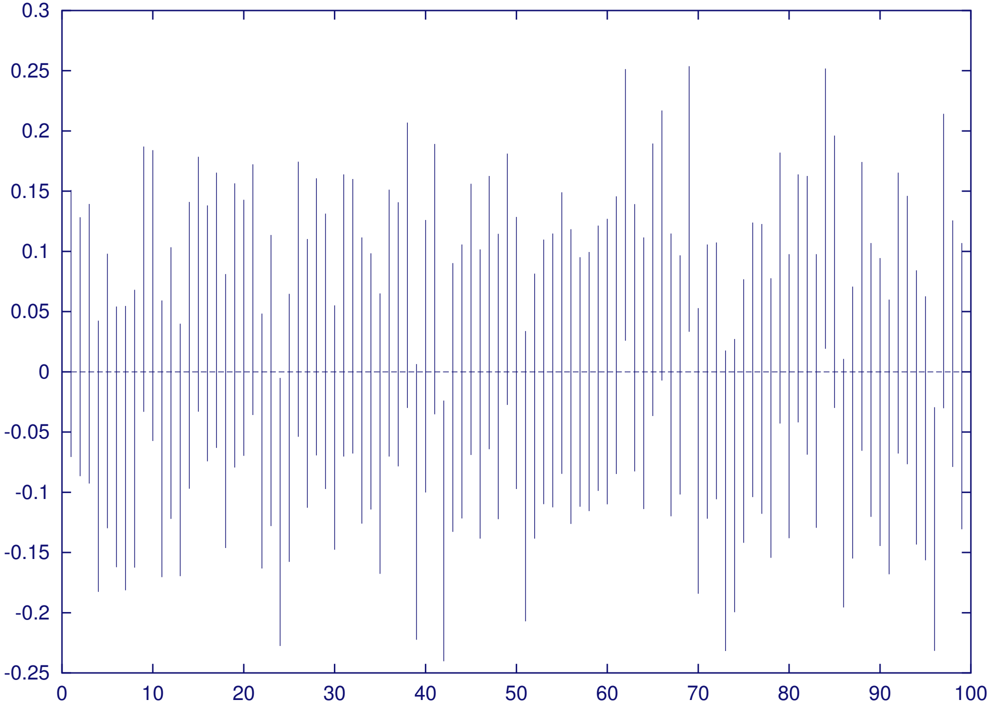
- For large n, an approximate $100(1-\alpha)\%$ confidence interval for μ is

$$\overline{X}(n) \pm z_{1-\frac{1}{2}\alpha}\frac{S(n)}{\sqrt{n}}$$

- To reduce the width of a confidence interval by a factor of x, about x^2 times as many observations are needed!

Interpretation

Beware: The confidence interval is random, not the mean μ!



100 confidence intervals for the mean of uniform random variable on $(-1, 1)$, where each interval is based on 100 observations

Joint random variables

If X and Y are discrete random variables, then

$$p(x, y) = P(X = x, Y = y)$$

is the joint probability mass function of X and Y, and

$$\begin{aligned} P_X(x) &= P(X = x) = \sum_y P(X = x, Y = y) \\ P_Y(y) &= P(Y = y) = \sum_x P(X = x, Y = y) \end{aligned}$$

are the marginal probability mass functions of X and Y.

Example:

Repeatedly draw from $1, \ldots, 10$. Let X be number of draws until 1 appears and Y until 10 appears. What is joint probability mass of X and Y?
What is the probability mass of $\min(X, Y)$?

Joint random variables

If X and Y are continuous random variables, then

$$P(X \leq a, Y \leq b) = \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} f(x,y)dxdy$$

is the joint probability distribution function of X and Y, where $f(x,y)$ is the joint density, satisfying

$$f(x,y) \geq 0, \quad \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f(x,y)dxdy = 1$$

The marginal densities of X and Y are

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy, \\ f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx. \end{aligned}$$

Joint random variables

- More generally, for every set B,

$$P((X,Y) \in B) = \int\int_B f(x,y)dxdy.$$

- Interpretation of joint density: for small $\Delta > 0$

$$P(x < X \le x + \Delta, y < Y \le y + \Delta) \approx f(x,y)\Delta^2$$

- Joint density follows by taking partial derivatives,

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} P(X \le x, Y \le y)$$

Joint random variables

Example:

X is distance to 0 and Y is angle (in radians) of a random point in a disk of radius r. Then:

- Joint distribution of X and Y is

$$P(X \leq x, Y \leq y) = \frac{x^2}{r^2}\frac{y}{2\pi}, \quad 0 \leq x \leq r, 0 \leq y \leq 2\pi$$

- Joint density of X and Y is

$$f(x, y) = \frac{2x}{r^2}\frac{1}{2\pi}, \quad 0 \leq x \leq r, 0 \leq y \leq 2\pi$$

Joint random variables

Example:

- Convolution: $Z = X + Y$ has density

$$f_Z(z) = \int_{-\infty}^{\infty} f(u, z-u)du.$$

- Let X be random on $(0, 1)$ and Y be random on $(0, X)$.
  What is the density of the area of the rectangle with sides X and Y?

Independent random variables

- The marginal densities of X and Y are

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

- X and Y are independent if

$$f(x,y) = f_X(x) f_Y(y) \quad \text{for all } x, y.$$

Independent random variables

Example:

- X is distance to 0 and Y is angle (in radians) of a random point in a disk of radius r. Then:

$$f_X(x) = \frac{2x}{r^2}, \quad 0 \le x \le r, \quad f_Y(y) = \frac{1}{2\pi}, 0 \le y \le 2\pi,$$

  and hence $f(x, y) = f_X(x) f_Y(y)$, so X and Y are independent

- Let X and Y be independent standard normal random variables. What is the distribution of $X^2 + Y^2$?

Independent random variables

Example:

Let $X_1, \ldots, X_n$ be independent exponentials with rates $\lambda_1, \ldots, \lambda_n$

- Probability that $X_i = \min\{X_1, \ldots, X_n\}$ is proportional to λ_i,

$$P(X_i = \min_j X_j) = P(X_i < \min_{j \neq i} X_j) = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}.$$

- Ordering of X_i and $\min_j X_j$ are independent,

$$\begin{aligned} P(X_i < \min_{j \neq i} X_j | \min_j X_j > t) &= P(X_i - t < \min_{j \neq i} X_j - t | \min_j X_j > t) \\ &= P(X_i < \min_{j \neq i} X_j). \end{aligned}$$

Size of minimum does not tell anything about whom of X_i is the minimum!

Substitution rule

Property: For any function g of the continuous X and Y with density $f(x, y)$,

$$E(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

(assuming that it exists).

Example:

Pick two points at random from the interval $(0, 1)$.
What is the mean distance between these two points?

Polar coordinates

The joint density of X and Y is $f(x, y)$ and let:

- R be distance of point (X, Y) to 0
- Φ be angle (in radians) of line from 0 to (X, Y) (between 0 and 2π)

Then the joint density of R and Φ is $f(rcos\phi, rsin\phi)r$

Example:

X and Y are independent standard normal random variables.
What is the joint density of R and Φ? And their marginal densities?

Covariance and correlation

- The covariance of X and Y is

$$\begin{aligned}\text{cov}(X,Y) &= E[(X-E(X))(Y-E(Y))] \\ &= E(XY)-E(X)E(Y)\end{aligned}$$

- If X and Y are independent, then $\text{cov}(X,Y)=0$.
- The correlation coefficient of X and Y is

$$\rho(X,Y)=\frac{\text{cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

- Property: $-1 \leq \rho(X,Y) \leq 1$.
- X and Y are uncorrelated if $\rho(X,Y)=0$
- Independent random variables are uncorrelated