

INTRODUCTION TO SPC

Ariane FERREIRA / Philippe CASTAGLIOLA

Ecole des Mines de Nantes & IRCCyN, Nantes, France

aproso@emn.fr

Process variability

- No product can be exactly the same.
- Why? There is a lot of sources of variability interfering with the process.
 - There are two main kinds of variability
 - **Common or Natural Causes** which are strongly linked with the process itself. These causes are considered as random and uncontrollable events.
 - **Special or Assignable Causes** are unpredictable events which make the process drift. These causes can be identified but are irregular and unstable.

Goals of SPC

- **SPC : Statistical Process Control.**

- **Goals**

1. To make the process able to produce within some specification limits, by centering it and by decreasing the natural variability due to the common causes. Main tool: **capability indices**.

2. To stabilize the process by identifying and removing the special causes.

Main tool: **control charts**.

Goals of SPC

- First *control charts* were developped during the 1930's by Walter A. Shewhart at Bell Labs.
- First capability indices were developped during the 1980's.
- This reflects two major steps in the Statistical Quality Control history
 - the *growing years* (1945-1975) with a product centered approach : controls only at the end of the line and a percentage of defects assumed "normal".
 - the *crisis years* (after 1975) with a process centered approach : the defects have to be anticipated → trend to the zero defect.

Why use SPC ?

- Approach suggested in ISO 900X.
- Data recording is a proof/protection against (internal/external) customers.
- Graphics are appealing and help people to continuously improve the process.
- Allow to understand better the process.
- The decreasing of the process natural variability reduces the number of defective items.
- ...

SPC data

| Samples | Measurements | | |
|----------|--------------|-----------|-----------|
| 1 | $X_{1,1}$ | $X_{1,2}$ | \dots |
| 2 | $X_{2,1}$ | $X_{2,2}$ | \dots |
| \vdots | \vdots | \vdots | \vdots |
| m | $X_{m,1}$ | $X_{m,2}$ | \dots |
| | | | $X_{m,n}$ |

- SPC data: m rows and n columns.
- $X_{j,k}$ is the measurement corresponding to the k th items of the j th sample.
- Normality must be checked (Anderson-Darling, Shapiro-Wilk).

Position characteristics

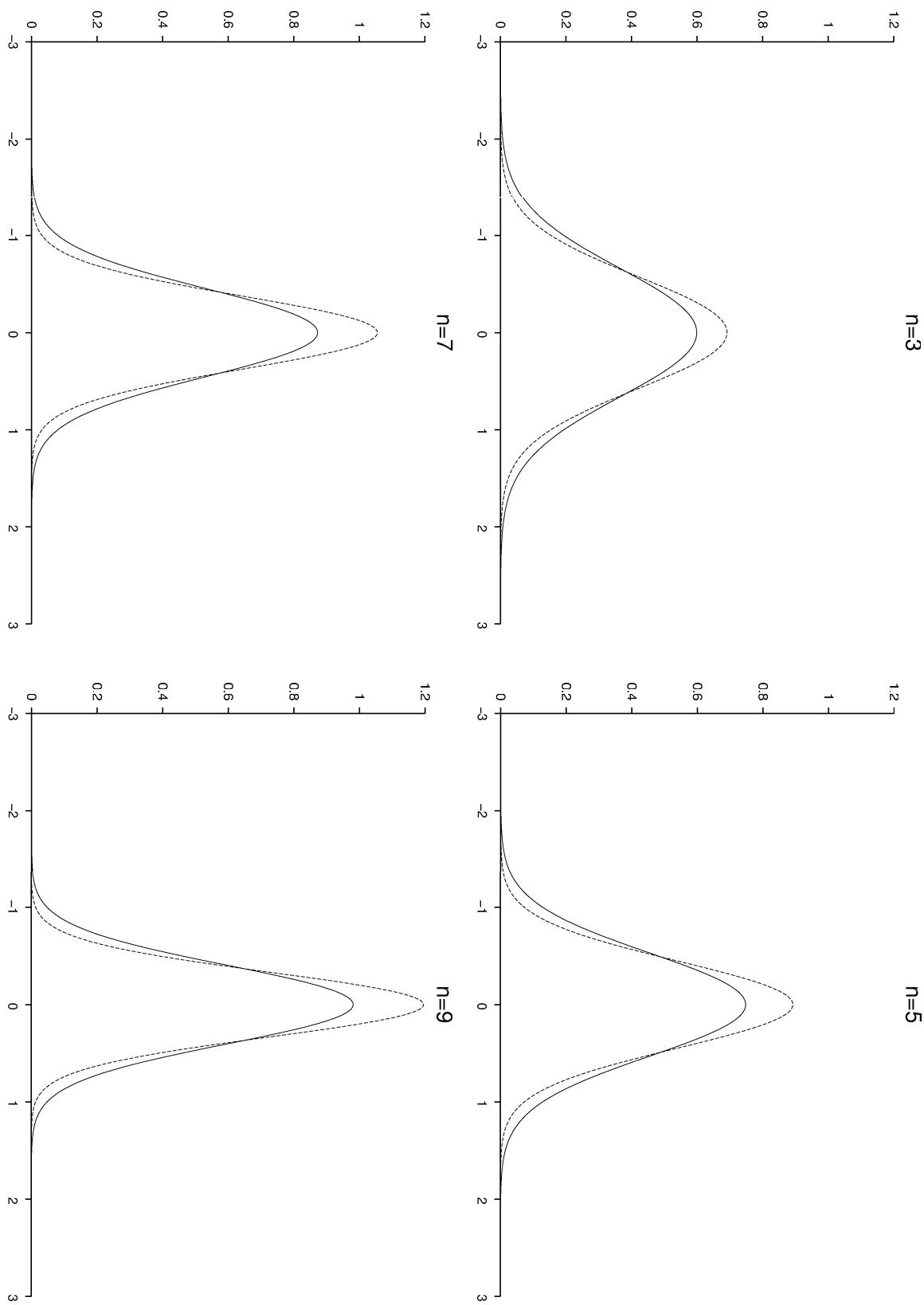
- The *mean* \bar{X}

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$$

- The *sample median* \tilde{X}

$$\tilde{X} = \begin{cases} X_{((n+1)/2)} & \text{if } n \text{ is odd} \\ \frac{X_{(n/2)} + X_{(n/2+1)}}{2} & \text{if } n \text{ is even} \end{cases}$$

- The sample median is more robust than the mean.
- If X_1, \dots, X_n are n iid normal (μ, σ) random variables, then \bar{X} and \tilde{X} are two unbiased estimators of μ .



| n | $\sigma(\bar{Z})$ | $\sigma(\tilde{Z})$ | $\gamma_2(\tilde{Z})$ |
|-----|-------------------|---------------------|-----------------------|
| 1 | 1.0000 | 1.0000 | 0.0000 |
| 3 | 0.5774 | 0.6698 | 0.0347 |
| 5 | 0.4472 | 0.5356 | 0.0320 |
| 7 | 0.3780 | 0.4587 | 0.0272 |
| 9 | 0.3333 | 0.4076 | 0.0233 |
| 11 | 0.3015 | 0.3704 | 0.0202 |
| 13 | 0.2774 | 0.3418 | 0.0178 |
| 15 | 0.2582 | 0.3189 | 0.0158 |
| 17 | 0.2425 | 0.3001 | 0.0143 |
| 19 | 0.2294 | 0.2842 | 0.0130 |
| 21 | 0.2182 | 0.2707 | 0.0119 |
| 23 | 0.2085 | 0.2589 | 0.0110 |
| 25 | 0.2000 | 0.2485 | 0.0102 |

Dispersion characteristics

- The standard-deviation $S > 0$

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

- The range $R > 0$

$$R = X_{(n)} - X_{(1)}$$

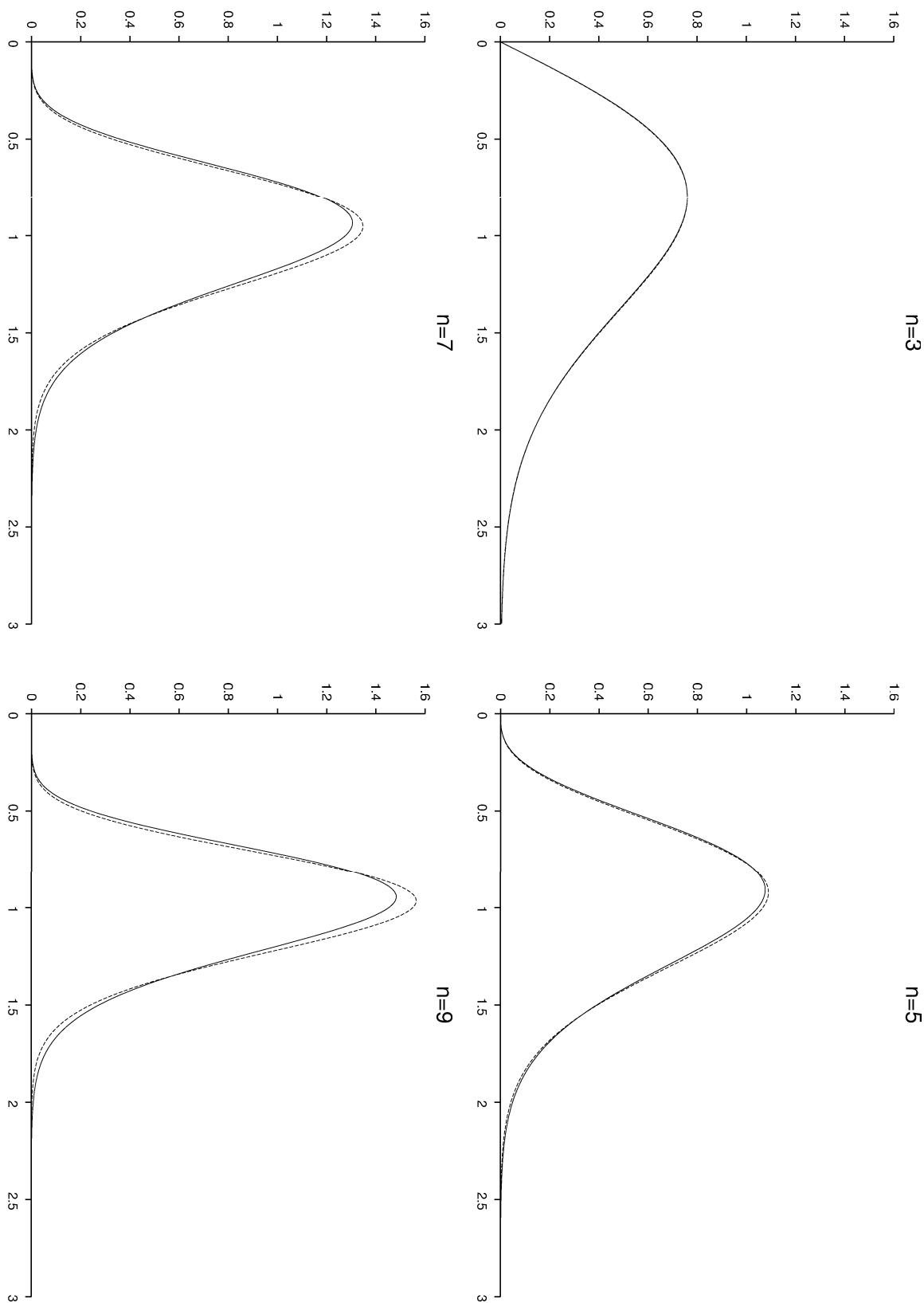
- Remark : S and R are not on the same scale.

Dispersion characteristics

- If X_1, \dots, X_n are n iid normal (μ, σ) random variables, then $S/K_S(n)$ and $R/K_R(n)$ are two unbiased estimators of σ , where

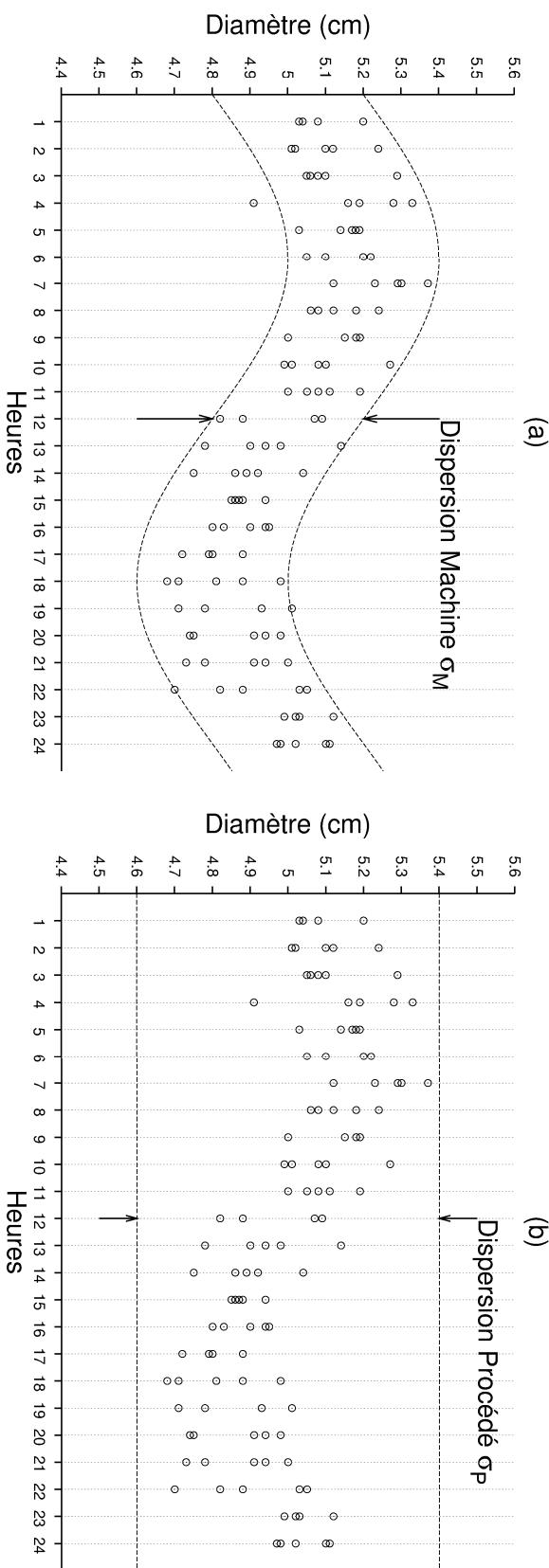
$$\begin{aligned} K_S(n) &= \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]} \sqrt{\frac{2}{n-1}} \simeq 1 - \frac{1}{4(n-1)} \\ K_R(n) &= 2 \int_0^{+\infty} 1 - \{\Phi(x)\}^n - \{1 - \Phi(x)\}^n dx \end{aligned}$$

| n | $K_S(n)$ | $K_S(n, -1)$ | $K_R(n)$ |
|-----|----------|--------------|----------|
| 2 | 0.7979 | | 1.1284 |
| 3 | 0.8862 | 1.7725 | 1.6926 |
| 4 | 0.9213 | 1.3820 | 2.0588 |
| 5 | 0.9400 | 1.2533 | 2.3259 |
| 6 | 0.9515 | 1.1894 | 2.5344 |
| 7 | 0.9594 | 1.1512 | 2.7044 |
| 8 | 0.9650 | 1.1259 | 2.8472 |
| 9 | 0.9693 | 1.1078 | 2.9700 |
| 10 | 0.9727 | 1.0942 | 3.0775 |
| 11 | 0.9754 | 1.0837 | 3.1729 |
| 12 | 0.9776 | 1.0753 | 3.2585 |
| 13 | 0.9794 | 1.0684 | 3.3360 |
| 14 | 0.9810 | 1.0627 | 3.4068 |
| 15 | 0.9823 | 1.0579 | 3.4718 |



Machine/Process variability

$$\sigma_M \leq \sigma_P$$



How to compute position/“dispersion” characteristics ?

- For the j th sample, one can compute
 - The mean \bar{X}_j .
 - The median \tilde{X}_j .
 - The standard-deviation S_j .
 - The range R_j .
- The process mean $\bar{X} = \frac{1}{m} \sum_{j=1}^m \bar{X}_j$
- The process median $\tilde{X} = \frac{1}{m} \sum_{j=1}^m \tilde{X}_j$

How to compute position/dispersion characteristics ?

- The machine standard-deviation $S_M = \frac{1}{m} \sum_{j=1}^m S_j$
- The machine range $R_M = \frac{1}{m} \sum_{j=1}^m R_j$
- The process standard-deviation $S_P = \sqrt{\frac{1}{mn - 1} \sum_{j=1}^m \sum_{k=1}^n (X_{j,k} - \bar{X})^2}$
- The process range $R_P = \max(X_{j,k}) - \min(X_{j,k})$

| | Diameters | | | | | \bar{X}_j | \tilde{X}_j | S_j | R_j |
|----|-----------|------|------|------|------|-------------|---------------|-------|-------|
| 01 | 4.87 | 5.26 | 4.97 | 5.12 | 5.03 | 5.05 | 5.03 | 0.148 | 0.39 |
| 02 | 5.15 | 4.92 | 5.07 | 5.03 | 5.04 | 5.04 | 5.04 | 0.083 | 0.23 |
| 03 | 5.03 | 4.93 | 5.11 | 5.10 | 5.16 | 5.07 | 5.10 | 0.089 | 0.23 |
| 04 | 5.17 | 5.14 | 5.20 | 5.29 | 5.06 | 5.17 | 5.17 | 0.084 | 0.23 |
| 05 | 5.23 | 5.07 | 5.08 | 5.23 | 5.14 | 5.15 | 5.14 | 0.078 | 0.16 |
| .. | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| 21 | 5.09 | 4.87 | 5.00 | 4.88 | 5.09 | 4.99 | 5.00 | 0.108 | 0.22 |
| 22 | 4.95 | 4.97 | 5.01 | 4.74 | 4.75 | 4.88 | 4.95 | 0.129 | 0.27 |
| 23 | 4.91 | 4.84 | 4.79 | 4.87 | 4.99 | 4.88 | 4.87 | 0.075 | 0.20 |
| 24 | 4.95 | 4.67 | 4.90 | 4.67 | 4.81 | 4.80 | 4.81 | 0.129 | 0.28 |
| | | | | | | \bar{X} | \tilde{X} | S_M | R_M |
| | | | | | | 5.09 | 5.09 | 0.095 | 0.23 |

Normality test for SPC data

- In order to test the normality of the *process*, then consider the m samples of size n as one sample of size mn .
- In order to test the normality of the *machine*, we have to remove the machine effect by making the following transformation

$$X_{j,k} \rightarrow X_{j,k} - \bar{X}_j$$

Estimation

- We assume $X_{j,k} \sim N(\mu, \sigma)$.

- Estimation of μ

$$\hat{\mu} = \bar{X} \quad \text{or} \quad \hat{\mu} = \tilde{X}$$

- Estimation of $\hat{\sigma}_M$

$$\hat{\sigma}_M = \frac{S_M}{K_S(n)} \quad \text{or} \quad \hat{\sigma}_M = \frac{R_M}{K_R(n)}$$

- Estimation of $\hat{\sigma}_P$

$$\hat{\sigma}_P = S_P$$