Statistical Process Control (SPC)

Nicholas Howe

Agenda

- What is SPC?
- Introduction to SPC Charts
- Examples
- Diagnostics

What is SPC?

- To detect out of control process:
  - Significant change in mean
  - Significant change in variance

- Examples:
  - Smartphone battery life
  - Particulate filtration efficiency

- Quality control
  - Six Sigma standards

https://www.digitaltrends.com/mobile/behind-scenes-samsung-designed-built-galaxy-s8-s8/
https://www.3m.com.sg/3M/en_SG/company-sg/all-3m-products/~/3M-Particulate-Respirator-8210-N95-160-Case/?N=5002385+3294780268&rt=rud
http://multimedia.3m.com/mws/media/3127030/masks-and-respirators-tri-fold-brochure-eng.pdf

Introduction to SPC Charts

Types of SPC Charts

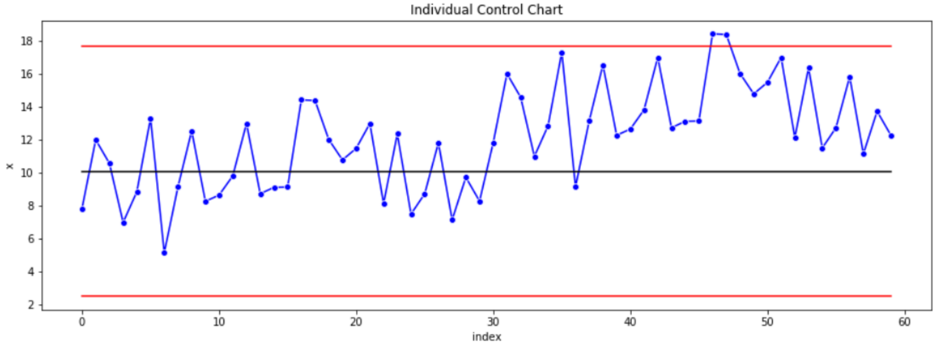
› Variable
  – Continuous data

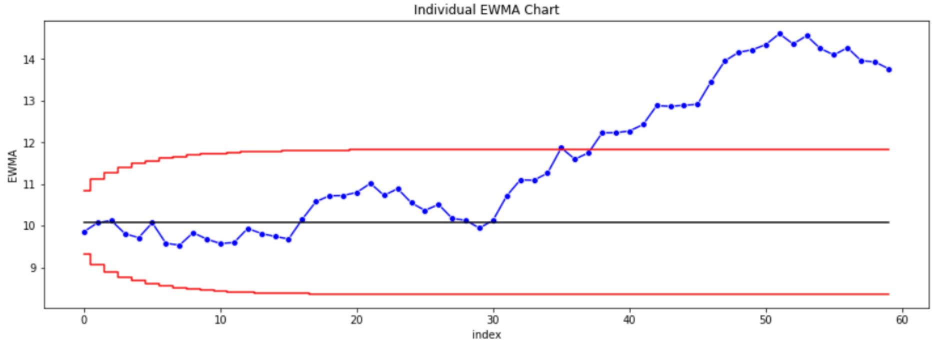
› Time Weighted
  – Increased sensitivity to out of control process

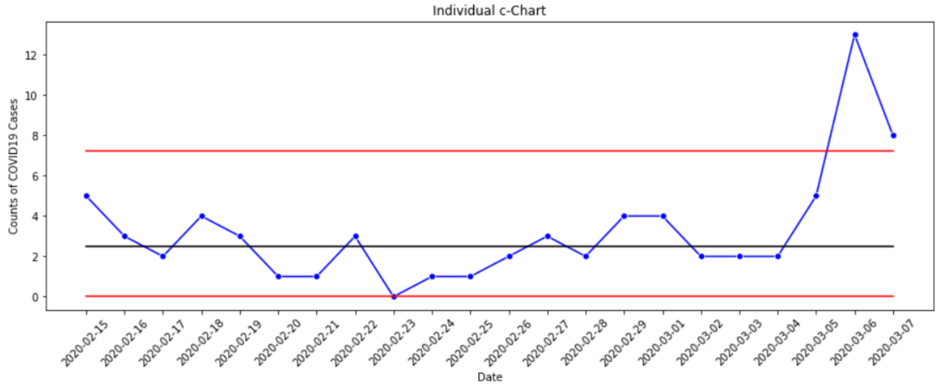
› Attribute
  – Counts data

Assumptions

- Independence
- Normally distributed



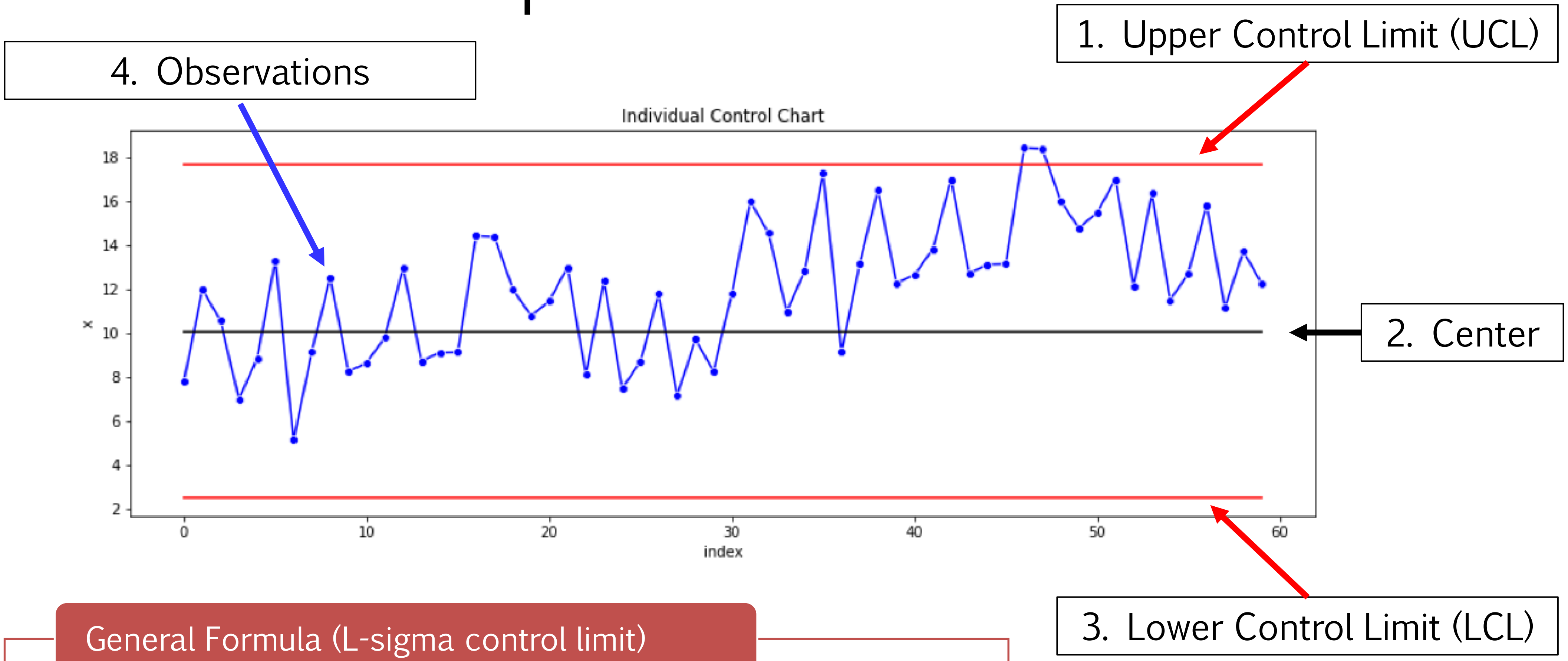




Summary of SPC charts

Variable	Time Weighted	Attribute
• $\bar{x}, s$	• $\boxed{\text{EWMA}}$	• p-chart
• $\bar{x}, R$	• Tabular Cusum	• np-chart
• $\boxed{x, R}$		• $\boxed{\text{c-chart}}$
• s		• u-chart
• s^2		
• R		

SPC Chart Components



General Formula (L-sigma control limit)

- $\text{UCL} = \mu_0 + L\hat{\sigma}$
- $\text{Center} = \mu_0$
- $\text{LCL} = \mu_0 - L\hat{\sigma}$

Setting Up SPC Charts

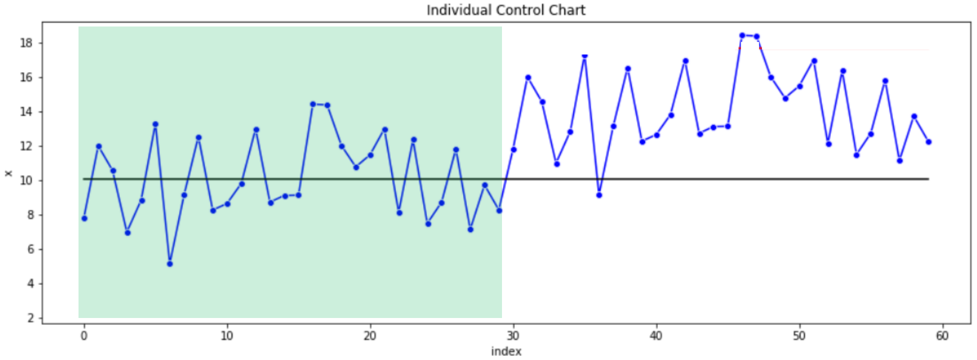
Phase I

- Get process under control or select a period where process is under control
- Estimate required parameters from in control data
- Construct SPC chart

Phase II

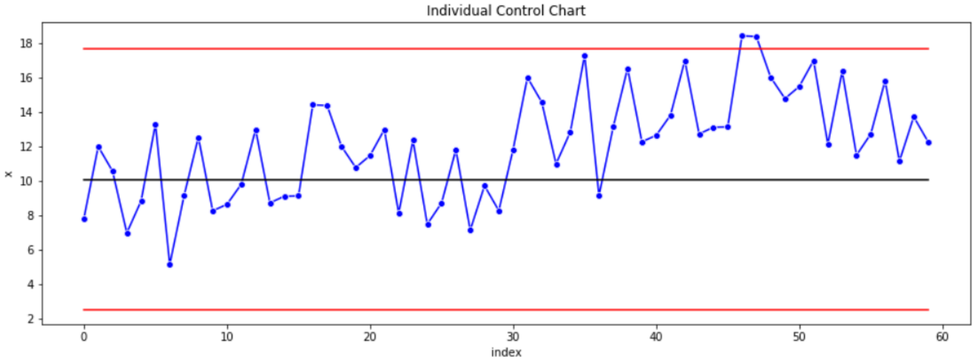
- Use SPC chart to monitor process
- Determine assignable cause when process signals out of control

Process in control



General Formula (L-sigma control limit)

- $\text{UCL} = \mu_0 + L\hat{\sigma}$
- $\text{Center} = \mu_0$
- $\text{LCL} = \mu_0 - L\hat{\sigma}$



Variable Control Charts: x, R Chart

- Individual observation chart (x)
- Parameter Estimation
  - $\mu_0 = \bar{x}$
  - $\hat{\sigma} = \overline{MR}/1.128$
- Moving Range:
  - $MR_i = |x_{i+1} - x_i|$
  - $\overline{MR} = \frac{1}{m-1} \sum_{i=1}^{m-1} MR_i$

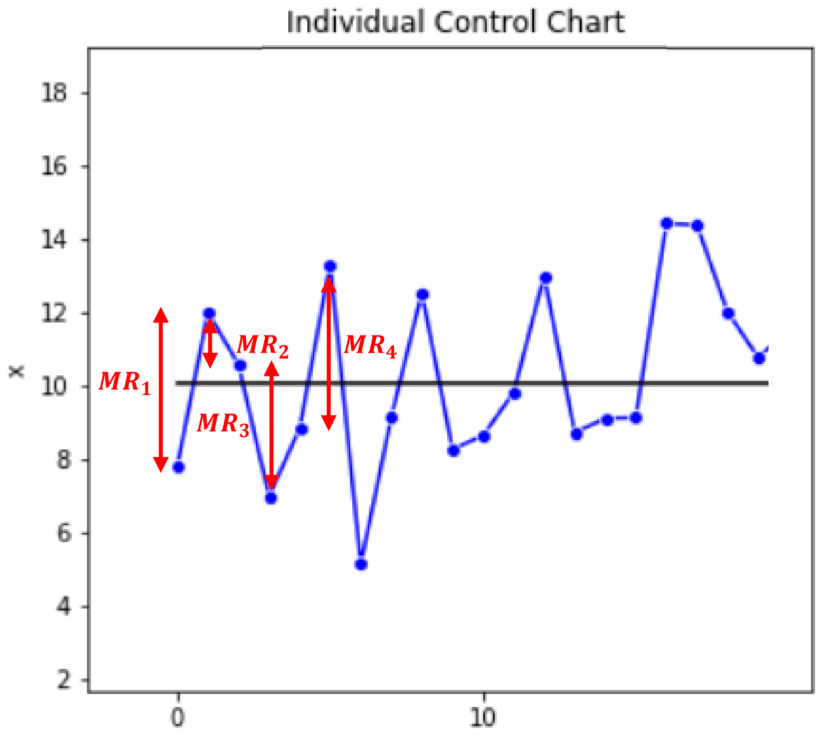


Chart Parameters (L-sigma control limit)

- $\text{UCL} = \mu_0 + L\hat{\sigma} = \mu_0 + L * \overline{MR}/1.128$
- $\text{Center} = \mu_0$
- $\text{LCL} = \mu_0 - L\hat{\sigma} = \mu_0 - L * \overline{MR}/1.128$

Note:
$W = R/\sigma$ is the relative range statistic.
When $X \sim N(\mu, \sigma^2)$ and $n = 2$, then
$E(R) = \sigma E(W) = 1.128\sigma$,
So estimate $\hat{\sigma}$ by $\overline{MR}/1.128$

Time Weighted Control Charts – EWMA

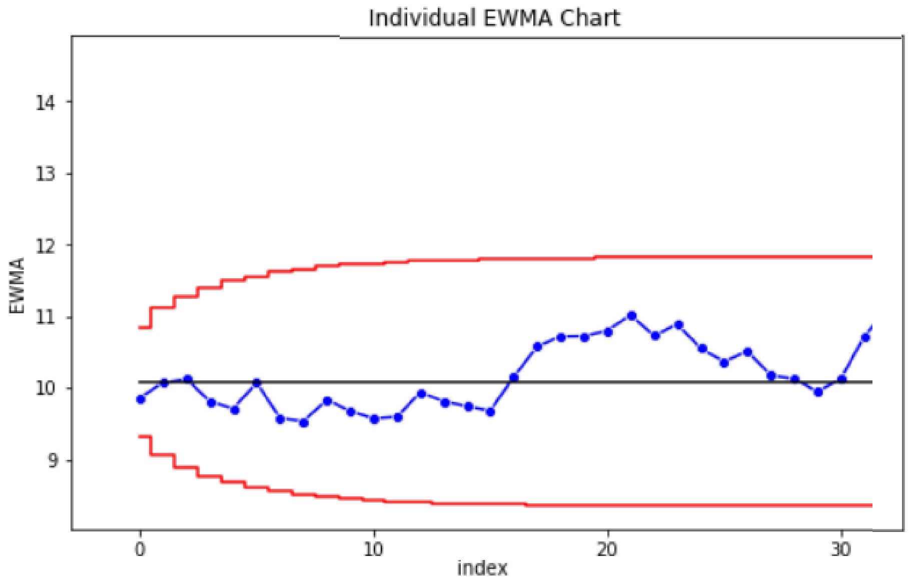
- Exponentially Weighted Moving Average (EWMA)
  - Better detection of small changes in mean
- Parameter Estimation
  - $\mu_0 = \bar{x}$
  - $\hat{\sigma} = \overline{MR}/1.128$

EWMA Formula

- $z_0 = \mu_0$
- $z_i = \lambda x_i + (1-\lambda) z_{i-1}, \quad 0 < \lambda < 1, \quad i = 1,2,\ldots$

Chart Parameters (L-sigma control limit)

- $\mathrm{UCL_i} = \mu_0 + L * \sigma \sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]}$
- $\mathrm{Center} = \mu_0$
- $\mathrm{LCL_i} = \mu_0 - L * \sigma \sqrt{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]}$



Note: Typically choose $\lambda = 0.1$

Attribute Control Charts – c-chart

- Useful for counts data
- Suppose $Y \sim Poisson(\mu_0)$, then $E(Y) = \mu_o$ and $Var(Y) = \mu_0$
- Parameter Estimation
  - $\mu_0 = \bar{x}$
- Non-zero lower limit
  - If $\mathrm{LCL} < 0$, set $\mathrm{LCL} = 0$

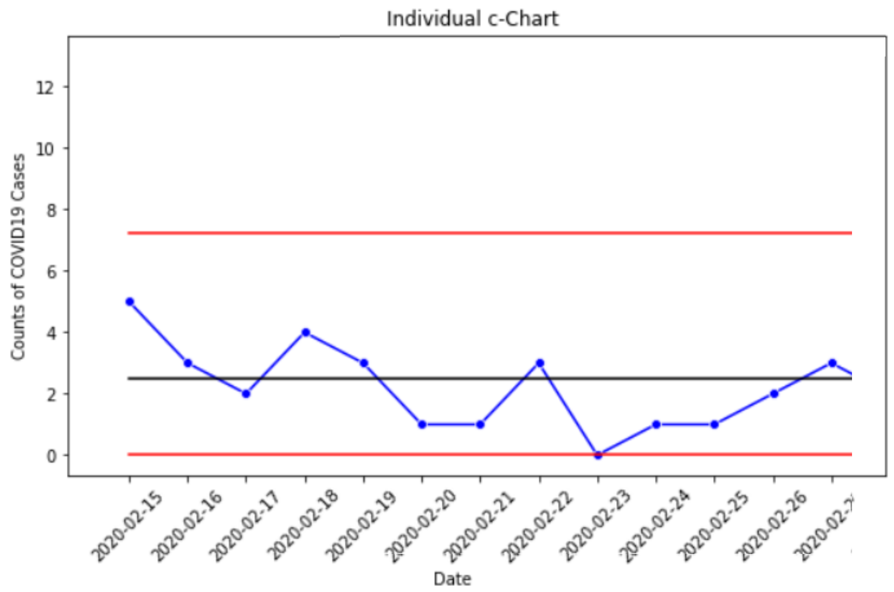


Chart Parameters (L-sigma control limit)

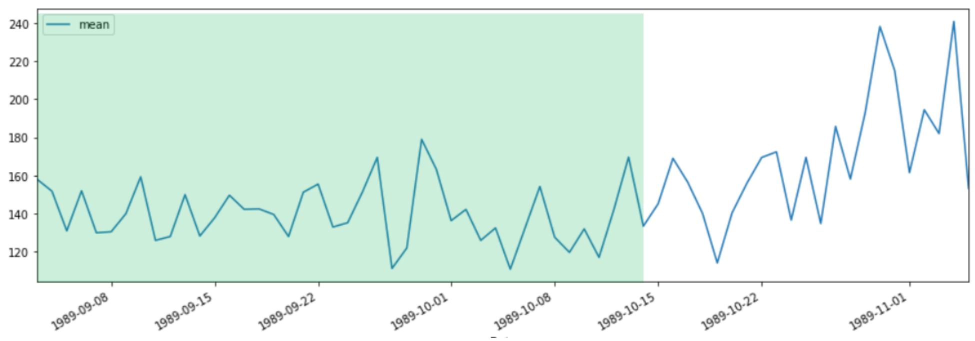
- $\mathrm{UCL} = \mu_0 + L\hat{\sigma} = \mu_0 + L * \sqrt{\mu_0}$
- $\mathrm{Center} = \mu_0$
- $\mathrm{LCL} = \mu_0 - L\hat{\sigma} = \mu_0 - L * \sqrt{\mu_0}$

Examples

Individual Blood Glucose Tracking (1/2)

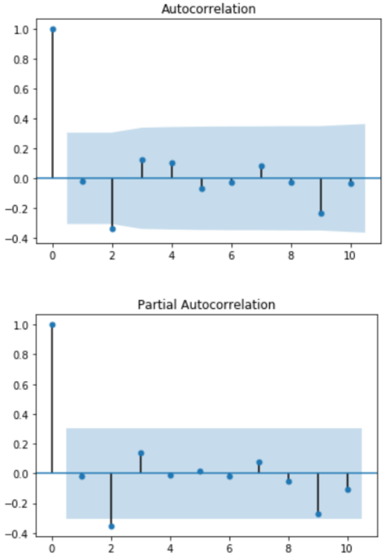
1. Select in-control data

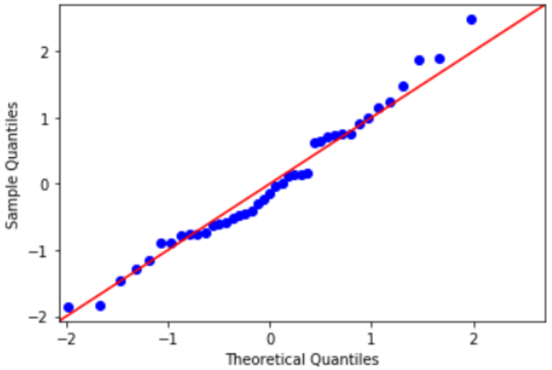
Process in control



- SPC Chart: Individual control chart
- Period selected: 1989-09-01 to 1989-10-14
- μ_0 is set to $\bar{x}$
- $\hat{\sigma}$ estimated by moving range

2. Check Assumptions

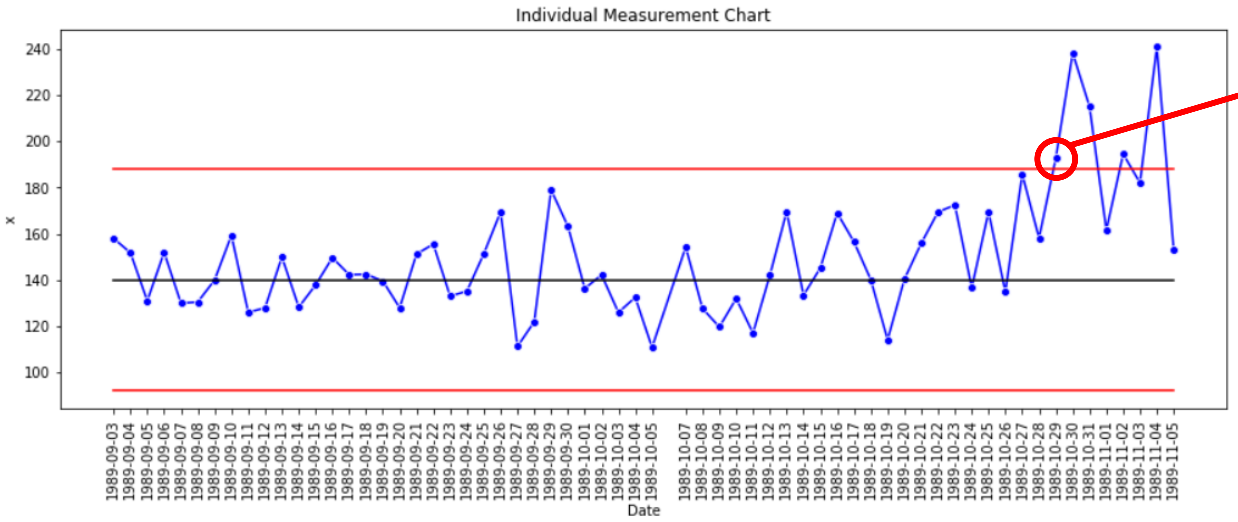




- Ljung–Box Test: No significant autocorrelation up to lag 10
- Jarque-Bera Test: Distribution is not significantly different from Gaussian

Individual Blood Glucose Tracking (2/2)

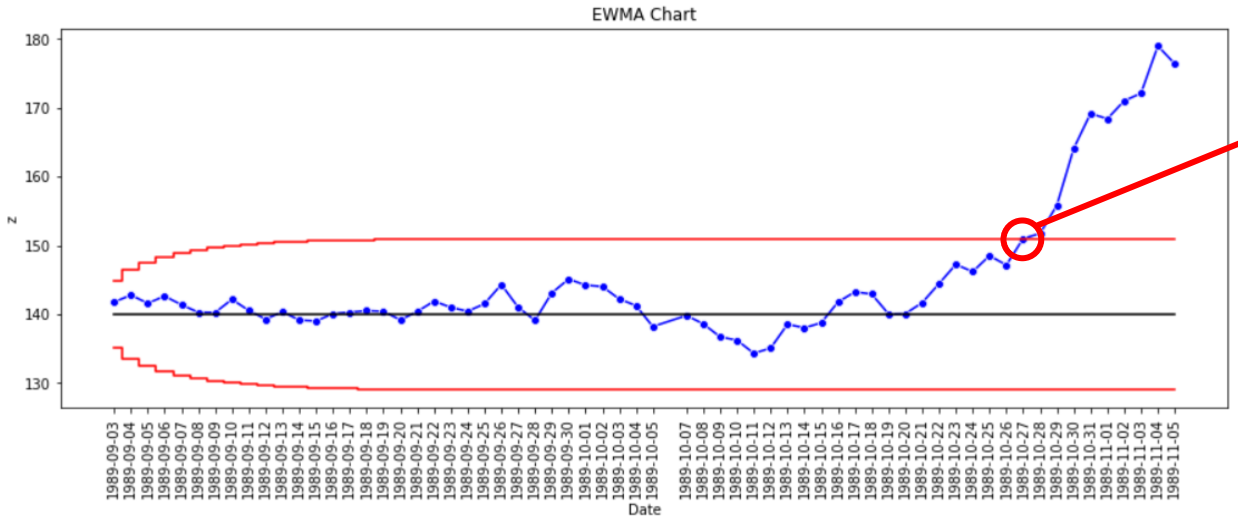
3. Construct Individual Control chart



Process first signals out of control on 1989-10-29

EWMA detects out of control mean 2 days earlier!

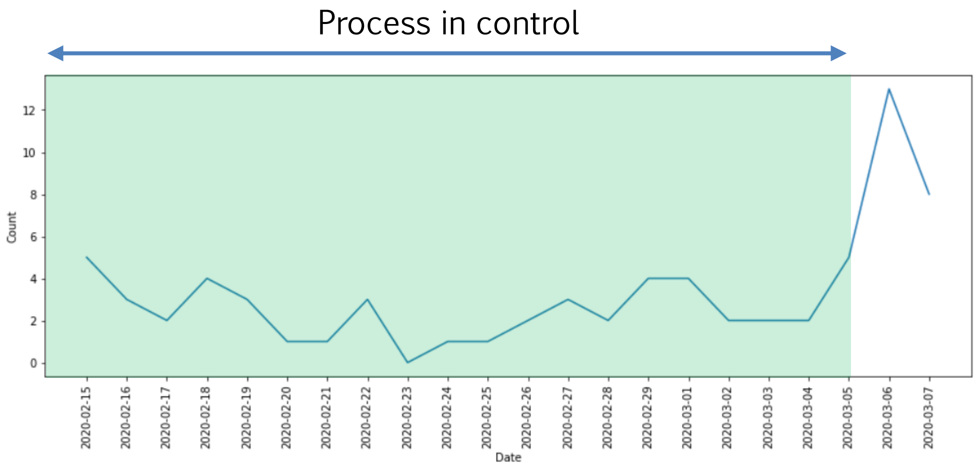
4. Construct EWMA Chart



Process first signals out of control on 1989-10-27

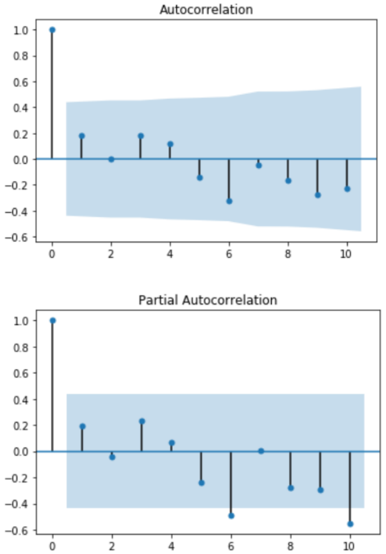
Singapore COVID-19 Cases (1/2)

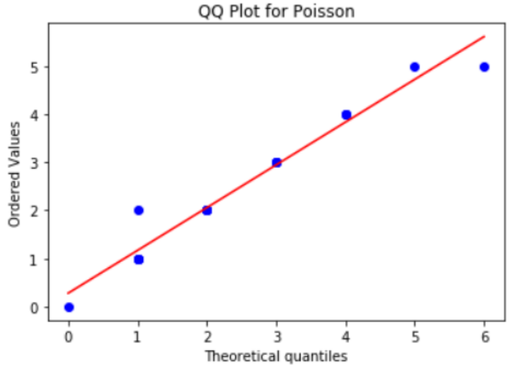
1. Select in-control data



- SPC Chart: c-chart
- Period selected: 2020-02-15 to 2020-03-05
- μ_0 is set to $\bar{x}$
- $\hat{\sigma}$ is set as $\sqrt{\bar{x}}$

2. Check Assumptions

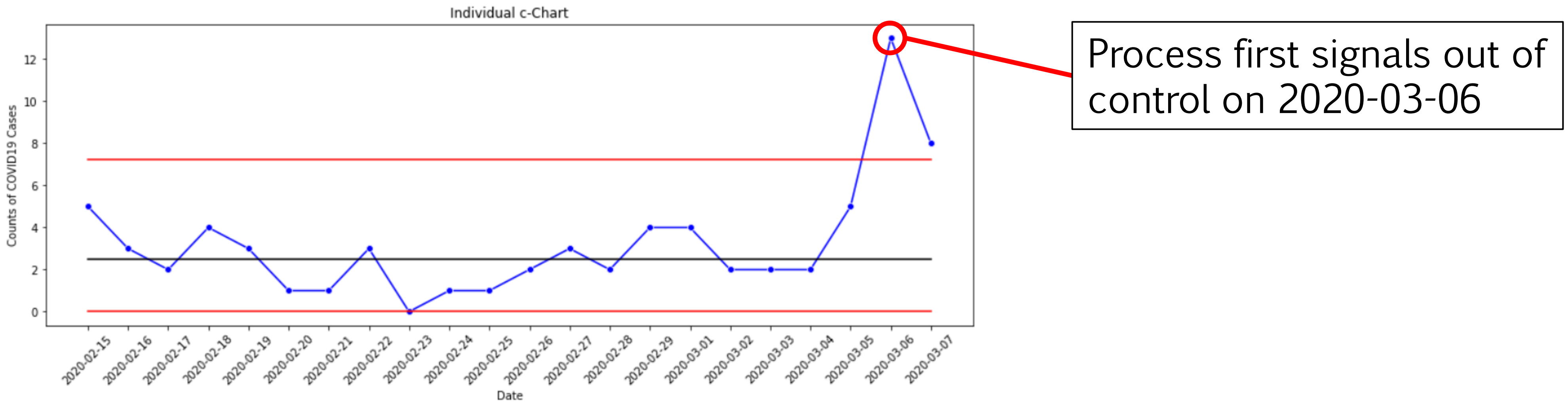




- Ljung–Box Test: No significant autocorrelation up to lag 10
- Chi-sq Goodness of fit Test: Distribution is not significantly different from Poisson

Singapore COVID-19 Cases (2/2)

3. Construct Individual Control chart



Diagnostics

Violations of Assumptions

- Violation of Normality
  - Skewed distributions
  - Heavy-tailed distributions

- Violation of Independence
  - Autocorrelated process

Properties of Normal Distribution

- Skewness = 0
- Kurtosis = 3

Skewed distributions (1/3)

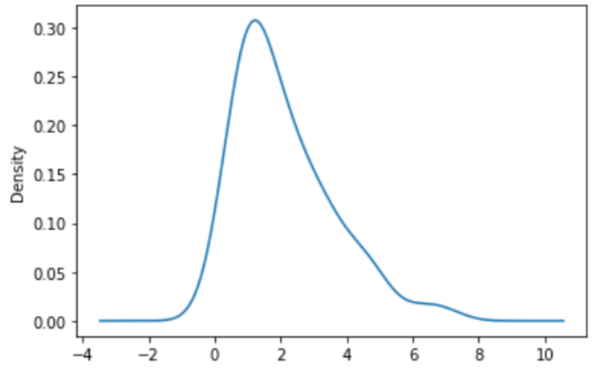
- **Results in false positives**
- Can be transformed into approximately Gaussian using the Yeo-Johnson Transformation
- The Yeo-Johnson transformation is defined as:

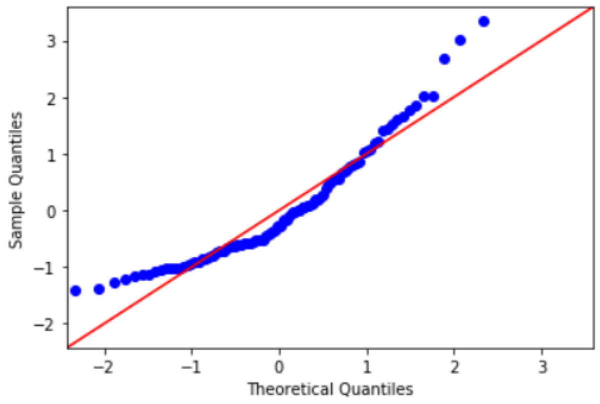
$$x_i^{(\lambda)} = \begin{cases} [(x_i+1)^\lambda - 1]/\lambda & \text{if } \lambda \neq 0, x_i \geq 0, \\ \ln(x_i+1) & \text{if } \lambda = 0, x_i \geq 0 \\ -[(-x_i+1)^{2-\lambda} - 1]/(2-\lambda) & \text{if } \lambda \neq 2, x_i < 0, \\ -\ln(-x_i+1) & \text{if } \lambda = 2, x_i < 0 \end{cases}$$

- Build individual control chart based on transformed data
- Reverse transform control limits and transformed data

Skewed distributions (2/3)
Gamma Distribution (Right Skewed)

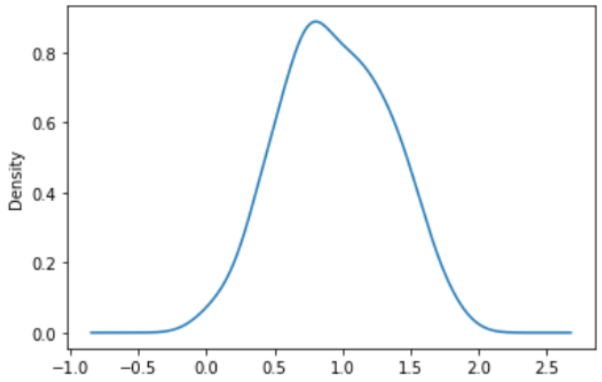
KDE and QQnorm for
Gamma Distribution

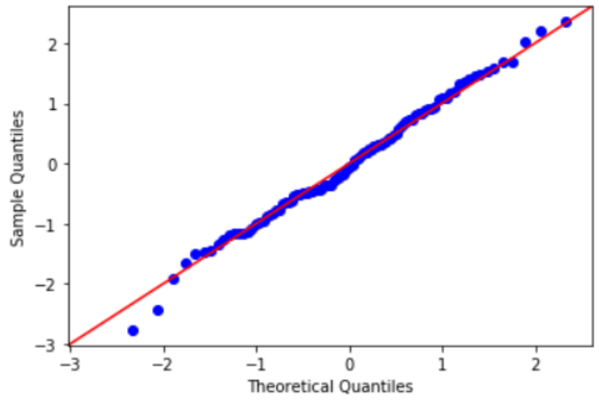




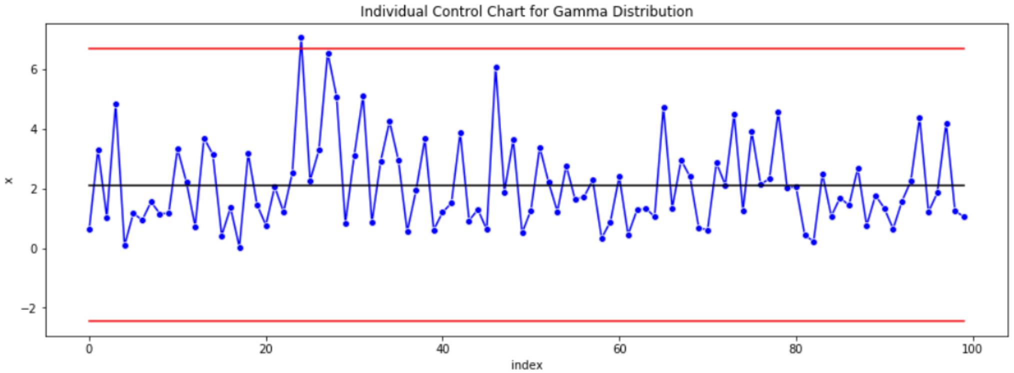
Yeo-Johnson
Transformation

KDE and QQnorm for
Transformed Distribution

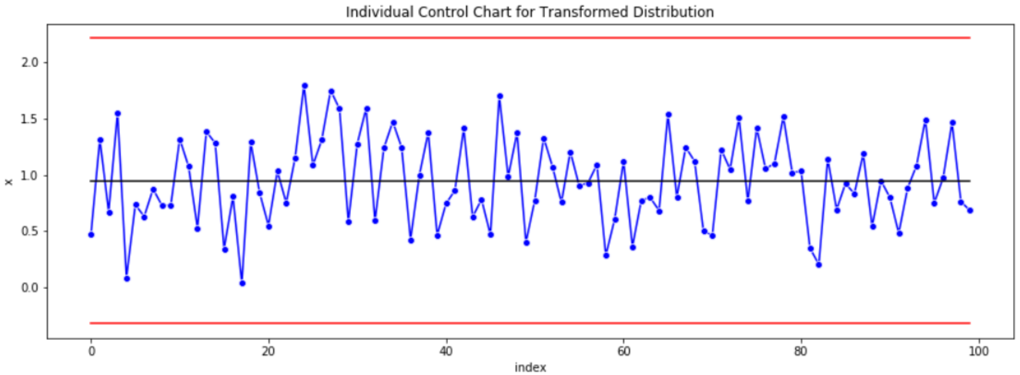




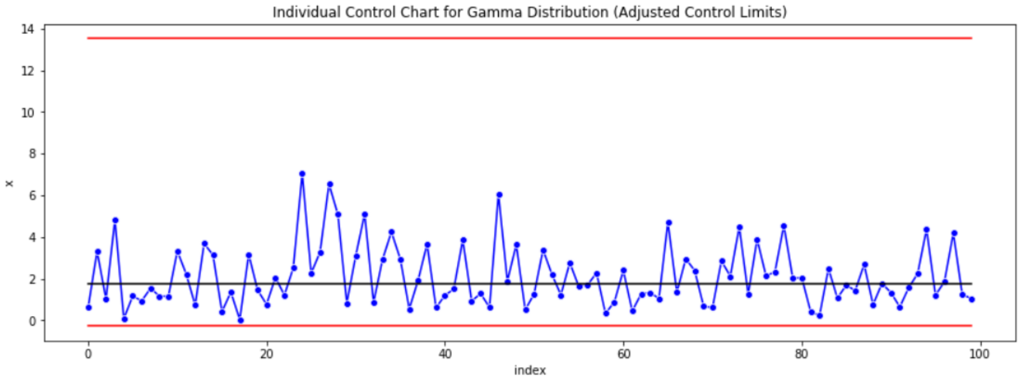
Skewed distributions (3/3)
Gamma Distribution (Right Skewed)



Yeo-Johnson Transformation



Reverse Transform Control Limits and transformed data

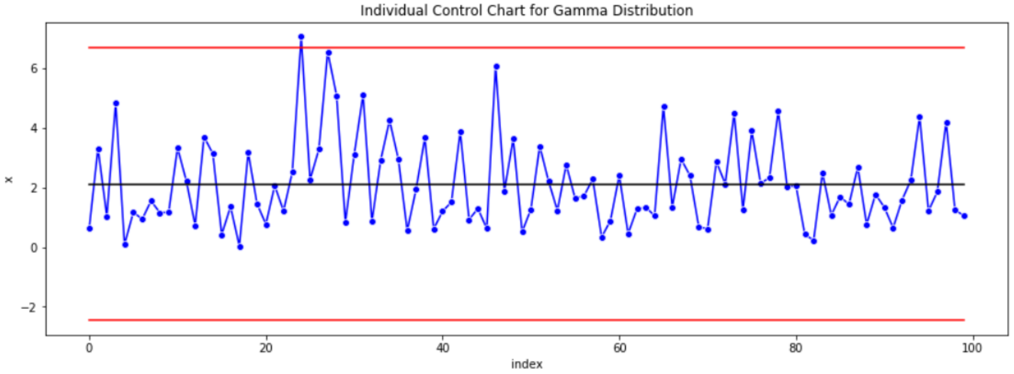


General Formula (L-sigma control limit)

- $\mathrm{UCL} = \mu_0 + L\hat{\sigma}$
- $\mathrm{Center} = \mu_0$
- $\mathrm{LCL} = \mu_0 - L\hat{\sigma}$

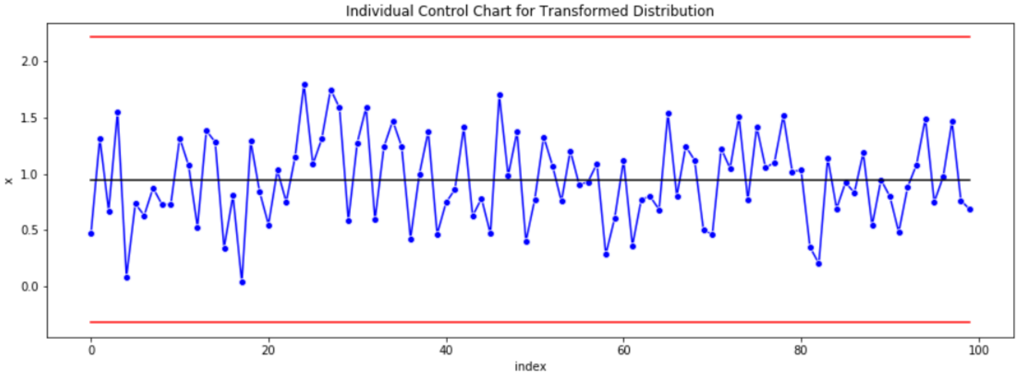
Skewed distributions (3/3)
Gamma Distribution (Right Skewed)

Construct SPC chart based on **original** distribution



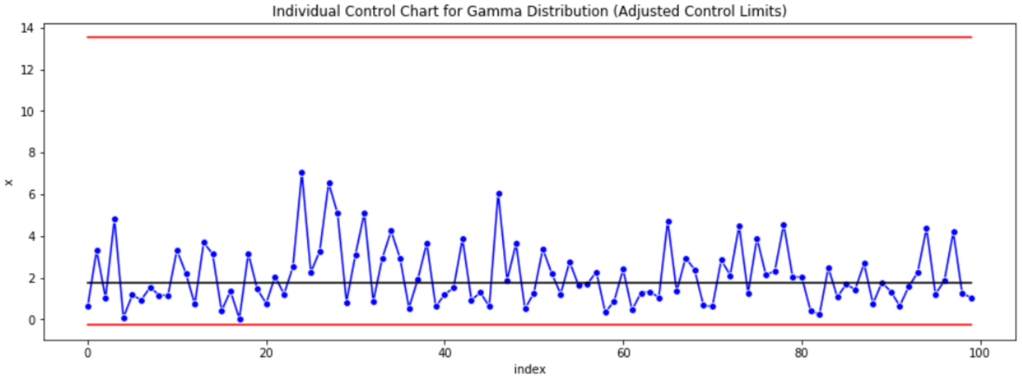
Yeo-Johnson Transformation

Construct SPC chart based on **transformed** distribution



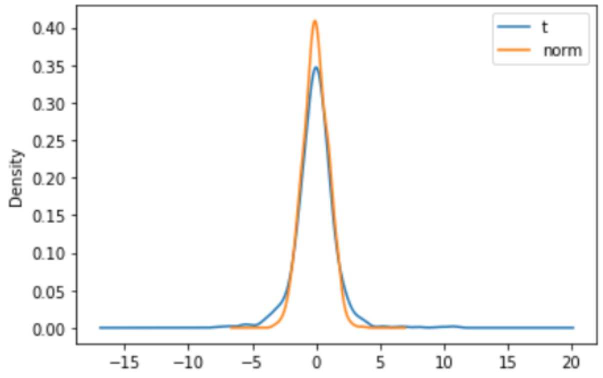
Back transform Control Limits and transformed observations

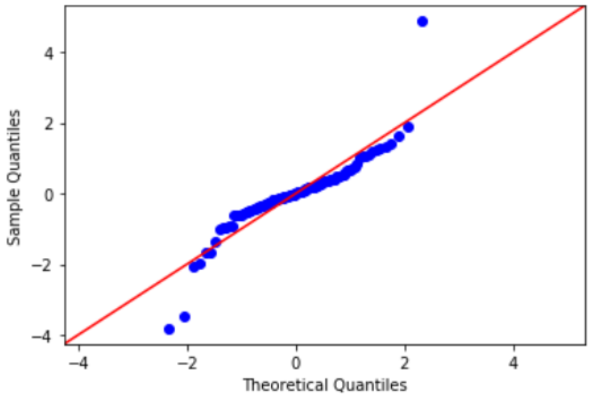
Original observations with back transformed control limits



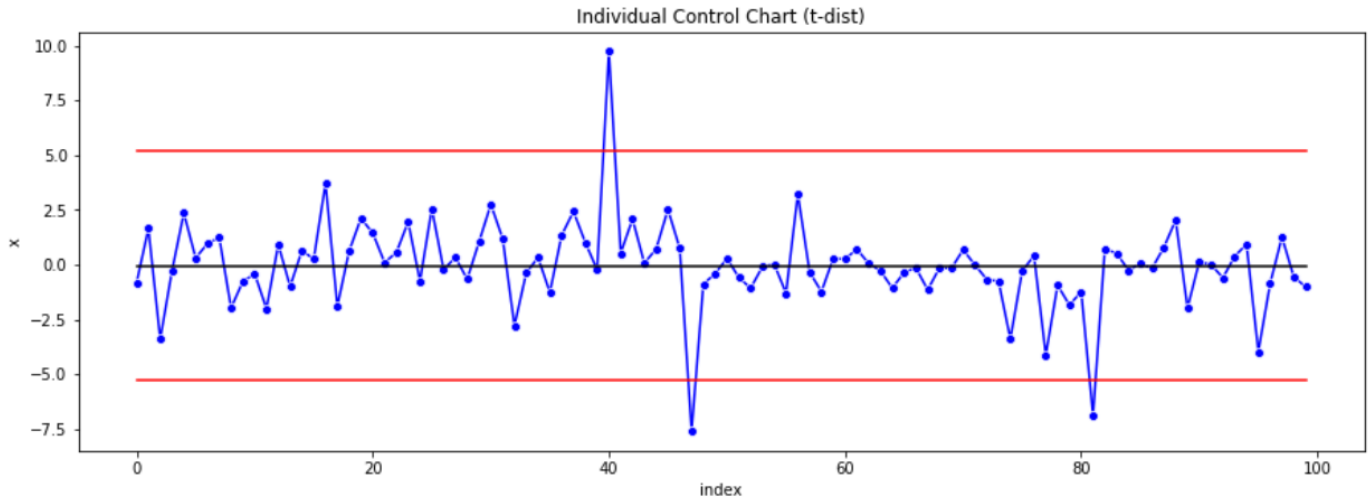
Heavy-tailed distributions (1/2) Student's t-distribution

KDE and QQnorm for t Distribution





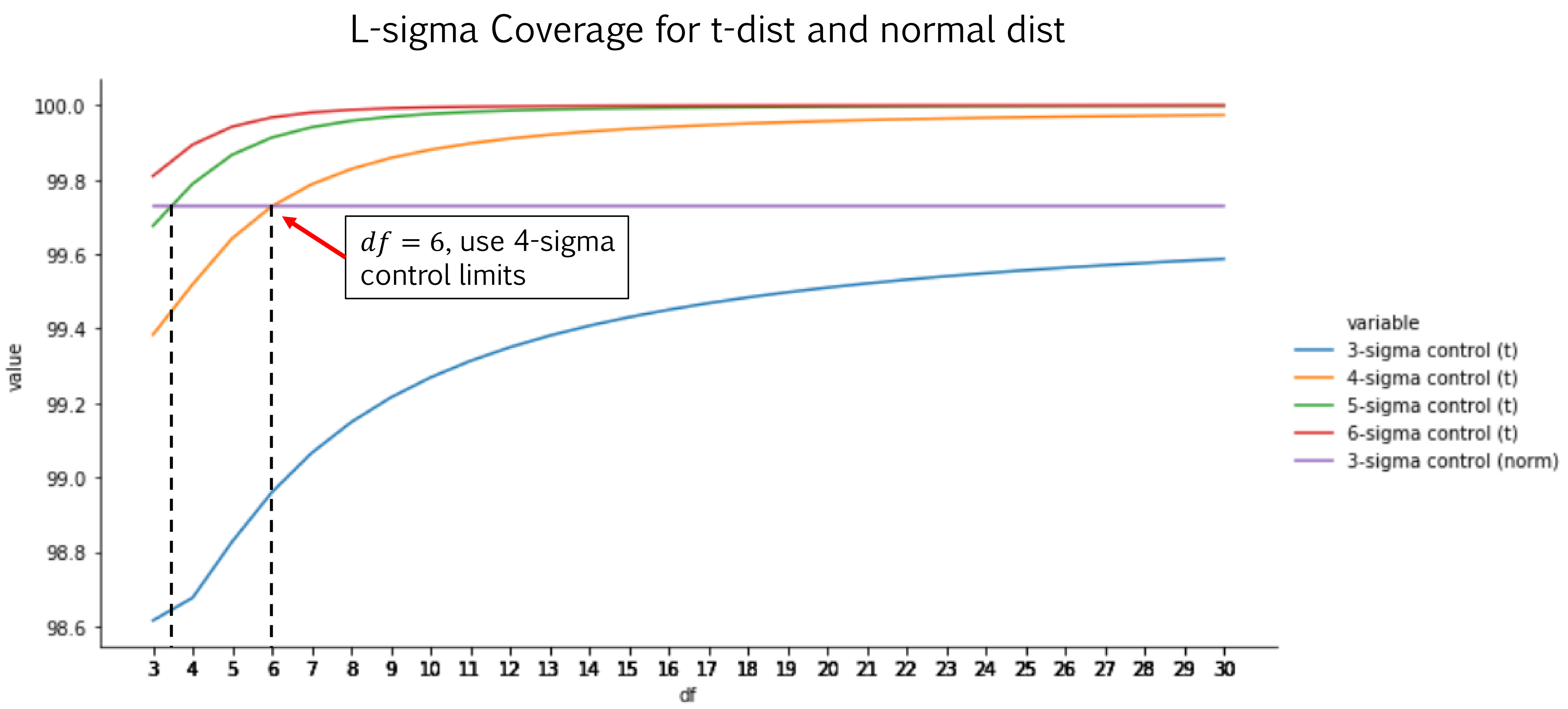
Individual Control Chart



Remedies

- Live with additional false positives
- Increase control limit

Heavy-tailed distributions (2/2) Student's t-distribution



Autocorrelated Process (1/2)
AR1 Simulation

- **Results in false positives**
- Fit an Auto-Regressive (AR) model to remove autocorrelation.
- In general, fit an AR1 model:

$$X_t = \beta_0 + \beta_1 X_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma^2)$$

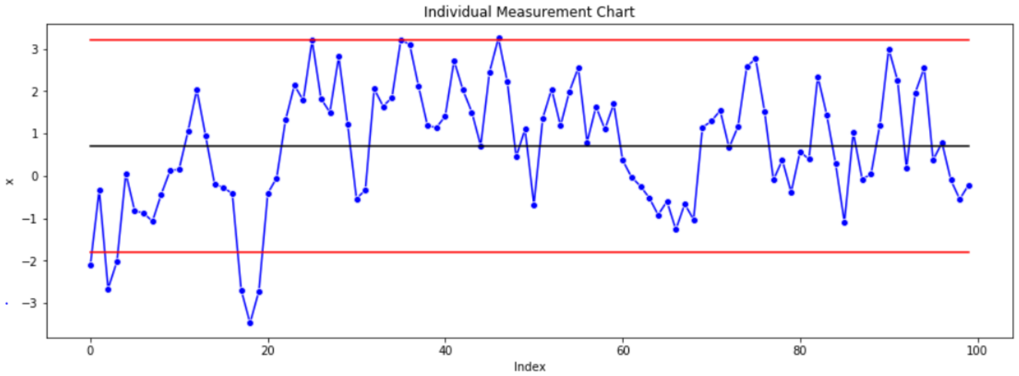
- Calculate residuals:

$$e_t = X_t - \hat{X}_t, \text{ where } \hat{X}_t = b_0 + b_1 X_{t-1}$$

- Build individual control chart based on residuals
- Reverse transform control limits and residuals by adding the predicted values to all the chart parameters
- E.g: $e_t + \hat{X}_t = X_t - \hat{X}_t + \hat{X}_t = X_t$

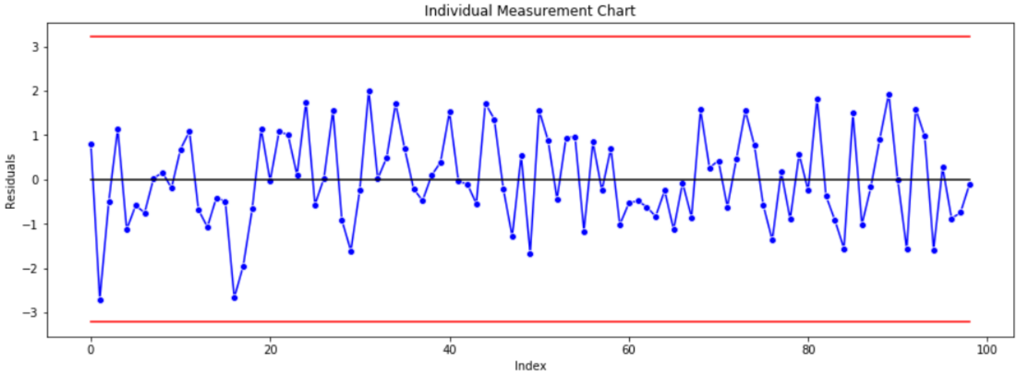
Autocorrelated Process (2/2)
AR1 Simulation

Construct SPC chart based on **original** distribution



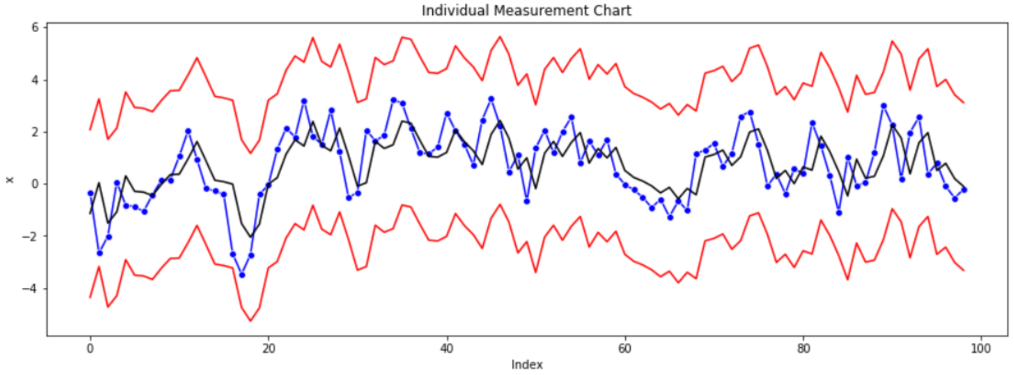
Fit AR1 model and calculate residuals

Construct SPC chart based on **transformed** distribution



Back transform Control Limits and residuals

Original observations with back transformed control limits



Thank You!

Appendix