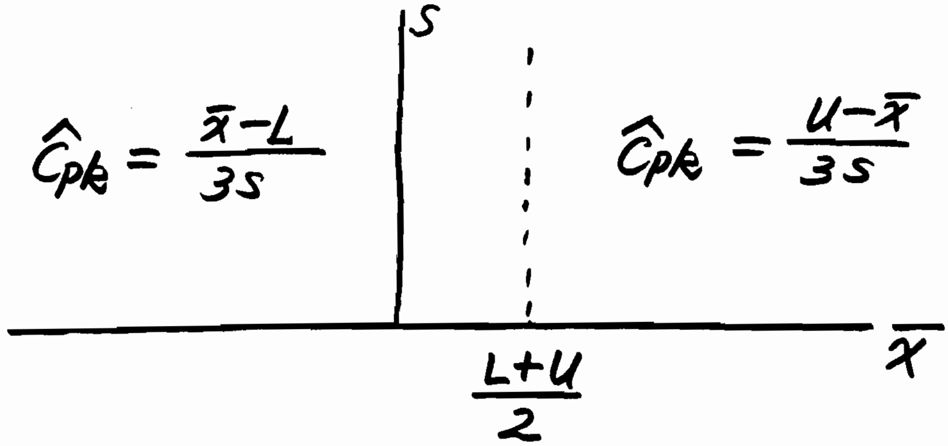
<u>Derivation</u>: $x_1, x_2, \cdots, x_n$ iid $N(\mu, \sigma^2)$




$$g(a,b) = \min\left\{ \frac{U-a}{3\sqrt{b}}, \frac{a-L}{3\sqrt{b}} \right\}$$

$$g(\bar{x}, s^2) = \hat{C}_{pk}$$

$\mathrm{Var}\, g(\bar{x}, s^2)$ can be approximated using $\Delta$ method.

Suppose $\mu > \frac{L+U}{2}$. Then

$$\mathrm{Var}\,\hat{C}_{pk} \approx \left(-\frac{1}{3\sigma}\right)^2 \mathrm{Var}(\bar{x}) + \left(\frac{U-\mu}{3}\right)^2 \left(-\frac{1}{2}(\sigma^2)^{-\frac{3}{2}}\right)^2 \mathrm{Var}(s^2)$$

now plug in $\mathrm{Var}\,\bar{x} = \frac{\sigma^2}{n}$

$$\mathrm{Var}(s^2) = \frac{2\sigma^4}{n-1}$$

and simplify to get $\frac{1}{9n} + \frac{C_{pk}^2}{2(n-1)}$.

Plug in $\hat{C}_{pk}$ for $C_{pk}$ and take the root and you have a standard error for $\hat{C}_{pk}$.

---

§5.3 Another way to characterize process output is through statements of where individual values are likely to be — prediction and tolerance intervals (PI's & TI's)

PI & TI problems are:

$X_1, X_2, \cdots, X_n$ iid $F \in \mathcal{F}$

some class of cont$^{\underline{s}}$ dsns

$X_{n+1}$

Make a random interval

$$(g_L(\underset{\sim}{X}), g_U(\underset{\sim}{X})) = I(\underset{\sim}{X})$$

from $\underset{\sim}{X} = (X_1, \cdots, X_n)$.

PI problem:

If $P_F[X_{n+1} \in I(\underset{\sim}{X})] \geq \gamma \quad \forall F \in \mathcal{F}$ (for all)

then I call $I(\underset{\sim}{X})$ a $\gamma$-level PI.

TI problem:

If $P_F[F(g_U(\underset{\sim}{X})) - F(g_L(\underset{\sim}{X})) \geq p] \geq \gamma$ $\forall F \in \mathcal{F}$, then I call $I(\underset{\sim}{X})$ a $\gamma$-level TI for a fraction $p$ of the underlying dsn $F$.

Simplest version of this is where

$\mathcal{F}$ = class of all cont's dsns

| Interval | Confidence Level Prediction Int. | Confidence Level as TI for fraction $p$ |
|---|---|---|
| $(\min X_i, \infty)$ | $\frac{n}{n+1}$ | $1-p^n$ |
| $(-\infty, \max X_i)$ | $\frac{n}{n+1}$ | $1-p^n$ |
| $(\min X_i, \max X_i)$ | $\frac{n-1}{n+1}$ | $1-p^n-n(1-p)\cdot p^{n-1}$ |

Why?

## One-Sided Intervals

As PI's:

$P[(\min X_i, \infty)$ doesn't cover $X_{n+1}]$

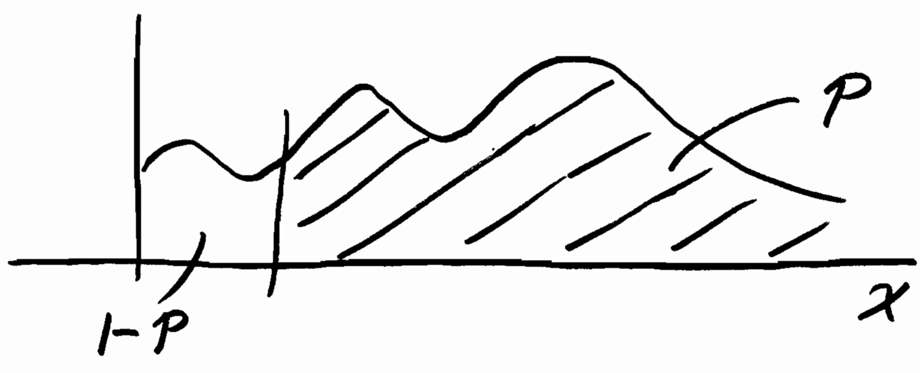
$= P[X_{n+1}$ is smallest of $X_1, X_2, \cdots, X_{n+1}]$

$= \frac{1}{n+1}$

$\Rightarrow$ confidence level is $\frac{n}{n+1}$ for prediction.

As TI's:

for a fraction $p$ of a dsn/process/population




$P[(\min X_i, \infty)$ fails to cover a fraction $p$ of the underlying dsn$]$

$= P[$all of $X_1, X_2, \cdots, X_n$ exceed the $1-p$ quantile of the underlying dsn$]$

$= p^n \Rightarrow$ confidence level is $1-p^n$.

## Two-Sided Interval

As PI:

$P[(\min X_i, \max X_i)$ doesn't cover $X_{n+1}]$

$= P[X_{n+1}$ is the smallest or largest of $X_1, X_2, \cdots, X_{n+1}]$

$= \frac{2}{n+1}$ by symmetry

$\Rightarrow$ confidence level is $\frac{n-1}{n+1}$.

As TI:

$P_F[F(\max X_i) - F(\min X_i) \geq p]$

$= P_F[\max F(X_i) - \min F(X_i) \geq p]$

and $X \sim F \Rightarrow F(X) \sim \text{Uniform}(0,1)$ $(U(0,1))$

—Why? roughly for $t \in (0,1)$

$$\begin{aligned} P[F(X) \leq t] &= P[X \leq F^{-1}(t)] \\ &= F(F^{-1}(t)) \\ &= t \end{aligned}$$

i.e., $F(X)$ is $U(0,1)$.

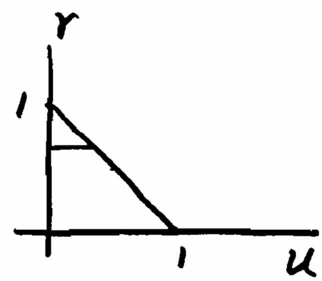
So we want

$P[\text{range of } n \text{ iid } U(0,1) \text{ r.v.'s} \geq p]$.

Parallel to material on page 2 of the notes: if $U, R$ are the minimum and range of $n$ iid $U(0,1)$ r.v.'s, their joint pdf is

$$h(u,r) = \begin{cases} n(n-1)r^{n-2} & \text{for } 0 \leq u \leq 1,\ 0 \leq r \leq 1, \\ & \text{and } 0 \leq u+r \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

For $r \in (0,1)$,




$$f(r) = \int_0^{1-r} n(n-1)r^{n-2}\,du$$

$$= n(n-1)r^{n-2}(1-r)$$

Hence $P[R \geq p] = \int_p^1 f(r)\,dr$

$$\vdots$$

$$= 1 - p^n - n(1-p)\cdot p^{n-1}$$

There is also a normal dsn technology for PI's and TI's.

<u>PI stuff</u> $X_1, \cdots, X_n, X_{n+1}$ iid $N(\mu, \sigma^2)$

Then $X_{n+1} - \bar{X} \sim N(0, \sigma^2(1+\frac{1}{n}))$

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

independent.

$$\Rightarrow T = \frac{X_{n+1} - \bar{X}}{S \cdot \sqrt{1+\frac{1}{n}}}$$

$$= \frac{(X_{n+1} - \bar{X}) / (\sigma \cdot \sqrt{1+\frac{1}{n}})}{\sqrt{\frac{(n-1)S^2}{\sigma^2} / (n-1)}}$$

$$= \frac{N(0,1) \text{ r.v.}}{\sqrt{\frac{\chi^2_{n-1} \text{ r.v.}}{n-1}}}$$ independent

$\sim t_{n-1}.$

So if $t^*$ and $t^{**}$ are such that

$$P[t^* < \text{a } t_{n-1} \text{ r.v.} < t^{**}] = \gamma$$

$$P[\underbrace{t^* < T < t^{**}}] = \gamma$$

$$\bar{X} + t^* \cdot s \cdot \sqrt{1+\frac{1}{n}} < X_{n+1} < \bar{X} + t^{**} \cdot s \cdot \sqrt{1+\frac{1}{n}}$$

i.e., $\bar{X} \pm t \cdot s \cdot \sqrt{1+\frac{1}{n}}$

are normal theory prediction limits for $X_{n+1}$.

<u>TI stuff</u>

one-sided TI (one-sided tolerance limit)

— suppose I want a $\gamma$ level <u>upper</u> tolerance bound for a fraction $p$ of a normal distribution.

Try something of the form

$$\bar{X} + \tau s$$

I want

$$P[\bar{X} + \tau s \geq \overbrace{p \text{ quantile of the dsn}}^{Q_z(p) = \mu + z_p \cdot \sigma}] \geq \gamma$$

Note that

$$\bar{X} + \tau \cdot S \geq \mu + z_p \cdot \sigma$$

iff $\bar{X} - \mu - z_p \sigma \geq -\tau \cdot S$

iff $\dfrac{\bar{X} - \mu - z_p \sigma}{S} \geq -\tau$

iff $\dfrac{1}{\sqrt{n}} \cdot \dfrac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} - \sqrt{n}\, z_p}{\dfrac{S}{\sigma}} \geq -\tau$

iff $\dfrac{\text{a } N(0,1) \text{ r.v.} - \sqrt{n}\, z_p}{\sqrt{\dfrac{\text{a } \chi^2_{n-1} \text{ r.v.}}{n-1}}} \geq -\sqrt{n} \cdot \tau$

independent

$$\underbrace{\dfrac{\text{a } N(0,1) \text{ r.v.} - \sqrt{n}\, z_p}{\sqrt{\dfrac{\text{a } \chi^2_{n-1} \text{ r.v.}}{n-1}}}}_{}$$ independent

a noncentral $t_{n-1}$ r.v. with noncentrality parameter $-\sqrt{n}\, z_p$.

i.e., I want

$$-\sqrt{n}\, \tau = 1 - r \text{ quantile of the noncentral } t_{n-1}(-\sqrt{n}\, z_p).$$

i.e., I want

$$\tau = -\frac{1}{\sqrt{n}}\left(1-\gamma \text{ quantile of } t_{n-1}(-\sqrt{n}\,z_p)\right)$$

$$= \frac{1}{\sqrt{n}}\cdot\left(\gamma \text{ quantile of } t_{n-1}(\sqrt{n}\,z_p)\right)$$

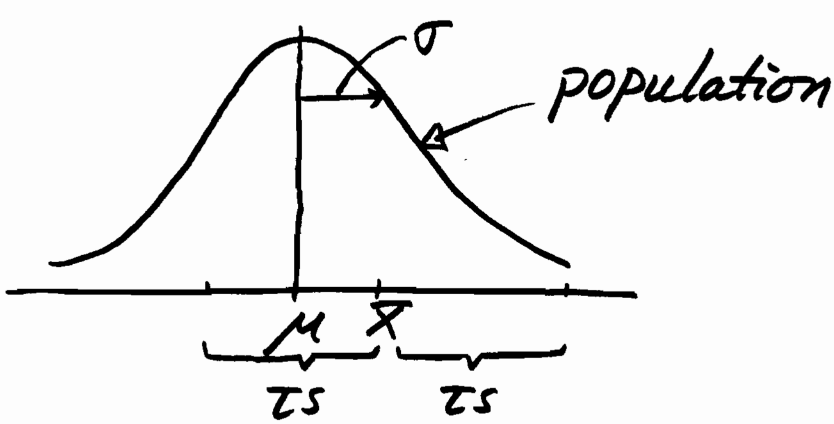
see page 523 of V&J, Table A.9b.

Notes: 1) There exist explicit approximations for $\tau$'s just considered (see problem 4.3 of the notes).

2) There exist two-sided one-sample TI's

$(\bar{x}-\tau\cdot s,\ \bar{x}+\tau\cdot s)$.

$$P_{\mu,\sigma}\left[\Phi\left(\frac{\bar{x}+\tau s-\mu}{\sigma}\right)-\Phi\left(\frac{\bar{x}-\tau s-\mu}{\sigma}\right)\geq p\right]\geq \gamma$$



This is the same as

$$P\left[\Phi\left(\frac{Z}{\sqrt{n}}+\tau\sqrt{\frac{W}{n-1}}\right)-\Phi\left(\frac{Z}{\sqrt{n}}-\tau\sqrt{\frac{W}{n-1}}\right)\geq p\right]\geq \gamma$$

where $\left.\begin{array}{l} Z \sim \text{std normal} \\ W \sim \chi^2_{n-1} \end{array}\right\}$ independent

— I must select $\tau$ to get the desired $\gamma$ (see page 522 of V&J, Table A.9a).

## §5.4. Probabilistic Tolerancing/Error Analysis

Technically this is nothing more than an application of propagation of error (simulation).

Application is to engineering design.

Idea is that in some contexts there are good deterministic models for how an output is a function of several inputs:

$$\underset{\text{some product quality variable}}{U} = g(\underbrace{X, Y, \cdots, Z}_{\text{properties of the product}})$$

First-order Taylor series analysis says for independent inputs

$$\mathrm{Var}\,U \approx \left(\frac{\partial g}{\partial x}\right)^2 \mathrm{Var}\,X + \left(\frac{\partial g}{\partial y}\right)^2 \mathrm{Var}\,Y + \cdots + \left(\frac{\partial g}{\partial z}\right)^2 \mathrm{Var}\,Z$$

where partials are evaluated at $(EX, EY, \cdots, EZ)$.

I can therefore hope to make $\mathrm{Var}\,U$ small by controlling $\mathrm{Var}\,X$, $\mathrm{Var}\,Y$, $\cdots$, $\mathrm{Var}\,Z$ or by making partials small (possibly by choice of $EX, EY, \cdots, EZ$).

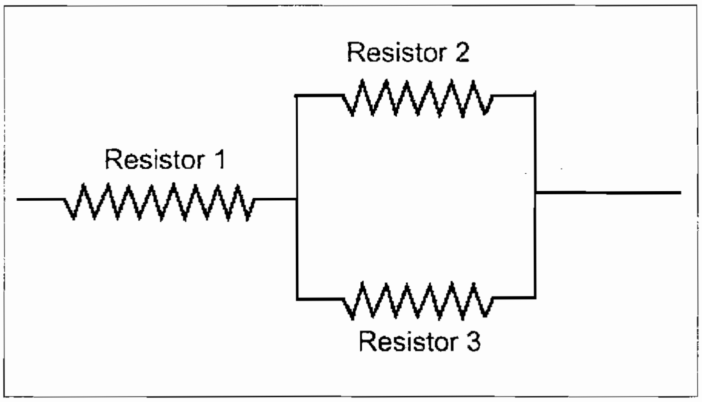
# Example 5.9

- Nice little tolerance stack-up problem

$$U = Y - X_1 - X_2 - X_3 - X_4$$
$$\sigma_U^2 = 1^2\sigma_Y^2 + (-1)^2\sigma_{X_1}^2 + (-1)^2\sigma_{X_2}^2 + (-1)^2\sigma_{X_3}^2 + (-1)^2\sigma_{X_4}^2$$

- $U$ was head space in a carton designed to hold 4 units of product

# Example 5.8

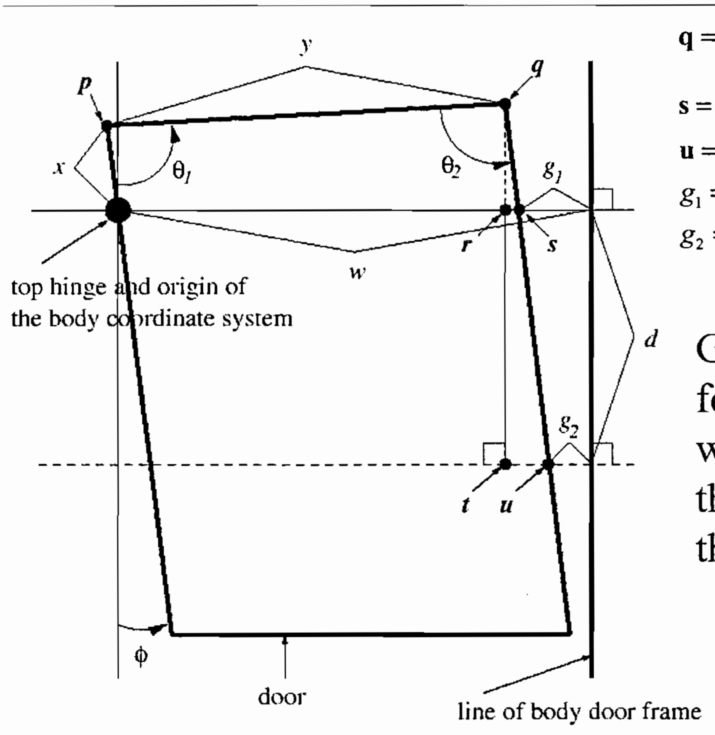



$$R = R_1 + \frac{R_2 R_3}{R_2 + R_3}$$

$$\mu_{R_1} = 100\Omega \text{ and } \mu_{R_2} = \mu_{R_3} = 200\Omega$$
$$\sigma_{R_1} = 2\Omega \text{ and } \sigma_{R_2} = \sigma_{R_3} = 4\Omega$$

What about $R$?

# Car Door Example




$$\mathbf{p} = (-x\sin\phi, x\cos\phi)$$

$$\mathbf{q} = \mathbf{p} + \left( y\cos\left(\phi + \left(\theta_1 - \frac{\pi}{2}\right)\right), y\sin\left(\phi + \left(\theta_1 - \frac{\pi}{2}\right)\right)\right)$$

$$\mathbf{s} = \left(q_1 + q_2 \tan(\phi + \theta_1 + \theta_2 - \pi), 0\right)$$

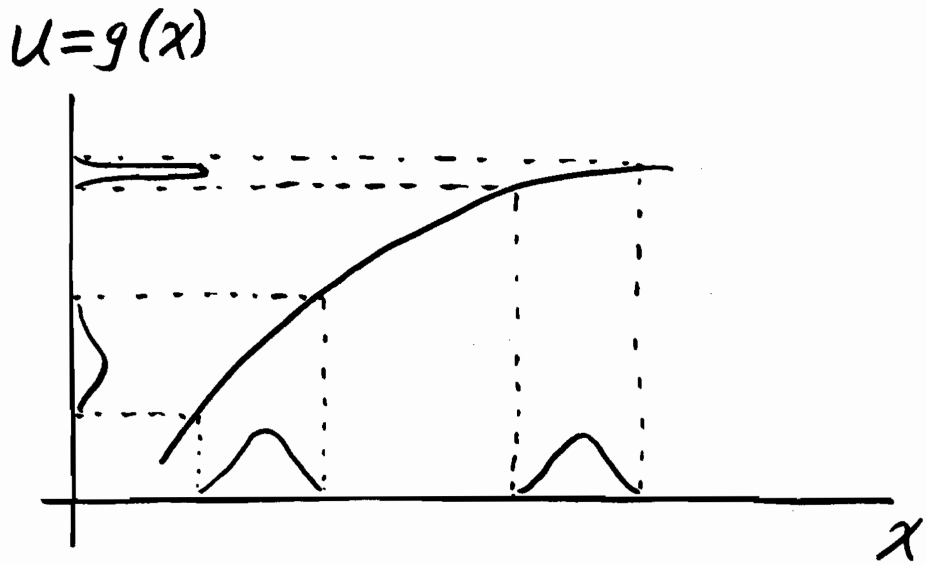
$$\mathbf{u} = \left(q_1 + (q_2 + d)\tan(\phi + \theta_1 + \theta_2 - \pi), -d\right)$$

$$g_1 = w - s_1$$

$$g_2 = w - u_1$$

Given nominals and "sigmas" for $x, y, w, \phi, \theta_1$ and $\theta_2$ what can I say about the the nominals and "sigmas" for the gaps, $g_1$ and $g_2$ ?

(Problem 4.12 of the Notes)




"Robust" ("Taguchi") product design

— see problems 4.11 and 4.12 of the notes for "examples."

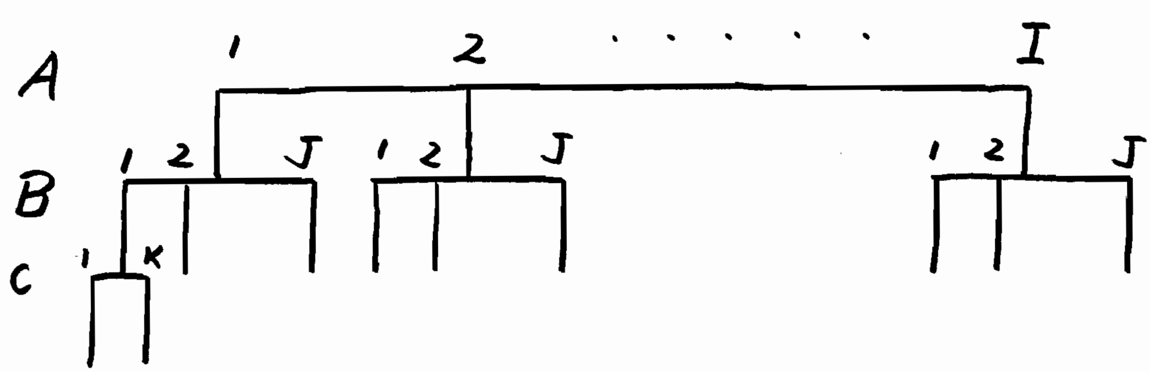
§5.5 of V&J (Ch. 4 of the notes)

## Variance Component Analysis in Hierarchical Contexts —

A common structure in industrial contexts is one where levels of A give rise to levels of B, which give rise to levels of C, etc. and one would like to "partition" variability encountered.

E.g.

| | | |
|---|---|---|
| A | Months | Heats |
| B | Weeks | Ingots |
| C | Days | Specimens |
| D | Shifts | Lab Determinations |




etc.

## Balanced Hierarchical Data

$y_{ijk}$ = observation for the kth level of C within jth level of B within the ith level of A

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + \varepsilon_{ijk}$$

for $\alpha$'s, $\beta$'s and $\varepsilon$'s independent

$$\alpha_i \sim N(0, \sigma_\alpha^2), \quad \beta_{ij} \sim N(0, \sigma_\beta^2), \quad \varepsilon_{ijk} \sim N(0, \sigma^2)$$

For balanced data, estimation of variance components is standard and based on ANOVA SS's and MS's.