§ 3 The stochastic project scheduling model

Now: random processing times X_j instead of fixed x_j

Assume that joint distribution P of $X = (X_1, \dots, X_n)$ is known

[Mostly, the X_i will be independent, but not necessarily]

$[G,P]$ is called a <u>stochastic project network</u>

For a (measurable) performance measure κ

always assumed in the sequel, holds for C_{max} and all the others considered so far

the distribution P_{κ^G} of $\kappa^G(.)$ is well defined

[Remember: $\kappa^G(x) = \underbrace{\kappa(\underbrace{ES_G(x) + x}_{\text{continuous}})}_{\text{measurable}}$]

THE UNDERESTIMATION ERROR OF DETERMINISTIC PLANNING

3.1 THEOREM: Let $[G,P]$ be a stochastic project network. Let $E(X) := (E(X_1), \dots, E(X_n))$ be the vector of the average processing times. If κ is convex, then

$$\kappa^G(E(X)) \le E(\kappa^G(\cdot))$$

i.e., the performance cost based on average processing times underestimates the expected performance cost

Proof: $$\kappa^G(x) = \underbrace{\kappa(\underbrace{ES_G(x)+x}_{\text{convex}})}_{\text{convex, since } \kappa \text{ is convex}} =: f(x_1,\dots,x_n)$$

$\Rightarrow$ Theorem follows with Jensen's inequality for convex functions

$$f(E(X_1),\dots,E(X_n)) \le E(f(X_1,\dots,X_n))$$

Expls for convex κ: C_{max}, $\sum w_j C_j$, $\sum w_j T_j$, T_{max}

<u>elementary proof for $\kappa = C_{max}$:</u>

Let $\mathcal{C} := \{C^1,\dots,C^m\}$ be the set of maximal chains of G

For C^i, let $Y_i := \sum_{j\in C^i} X_j$ be the (random) length of C^i

Then $C^G_{max} = \max_i Y_i$ and

$$C^G_{max}(E(X)) = \max_i \sum_{j\in C^i} E(X_j) = \max_i E\Big(\sum_{j\in C^i} X_j\Big)$$

$$= \max_i E(Y_i) = E(Y_{i_0})$$

assume maximum attained for $i=i_0$

$$\le E(\max_i Y_i) \quad \text{since } Y_{i_0} \le \max Y_i \text{ (as functions)}$$

$$= E(C^G_{max}) \quad \square$$

REMARKS:

(1) Equality holds for C_{max} iff one chain is critical with probability 1

(2) Error can get arbitrarily large

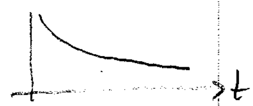
- with growing n
- with fixed n and growing variance of the X_i

Expl: ① ② ⋮ ⓝ

G_n = n-element antichain

Every $\underbrace{X_i \sim \exp(\lambda)}_{\text{exponential distribution}}$ with $\lambda = 1$, independent

with density



$f(t) = \lambda e^{-\lambda(t)}$ on $[0, \infty[$

$$E(X_i) = \frac{1}{\lambda}$$

Then $E(C_{max}^{G_n})$ = E(until first completion) + E(remainder)

$$= E(\min_j X_j) + E(C_{max}^{G_{n-1}})$$

↑ memoryless property of exponential distribution

$$= \frac{1}{\lambda_1 + \dots + \lambda_n} + E(C_{max}^{G_{n-1}})$$

↑ min of exp. distr. $X_1 \sim \exp(\lambda_1), \dots, X_n \sim \exp(\lambda_n)$ is $\exp(\lambda_1 + \dots + \lambda_n)$

$$= \frac{1}{n} + E(C_{max}^{G_{n-1}})$$

induction $\Rightarrow$ $E(C_{max}^{G_n}) = \sum_{i=1}^{n} \frac{1}{i}$ (harmonic series)

$\sim \log n$ + Eulerian constant .577

On the other hand, $C_{max}^{G_n}(E(X)) = 1$

$\Rightarrow$ absolute error, relative error $\to \infty$ $\square$

Summary:

- stochastic influences are important, need to be considered
- has led to PERT method of NASA

  $\uparrow$

  bad, considers only distribution of Y_{i_0}

  with $E(Y_{i_0}) = \max_i E(Y_i)$