A Dependent Project Evaluation and Review Technique, A Bayesian Network Approach

"Presentation Short Course: Beyond Beta and Applications"

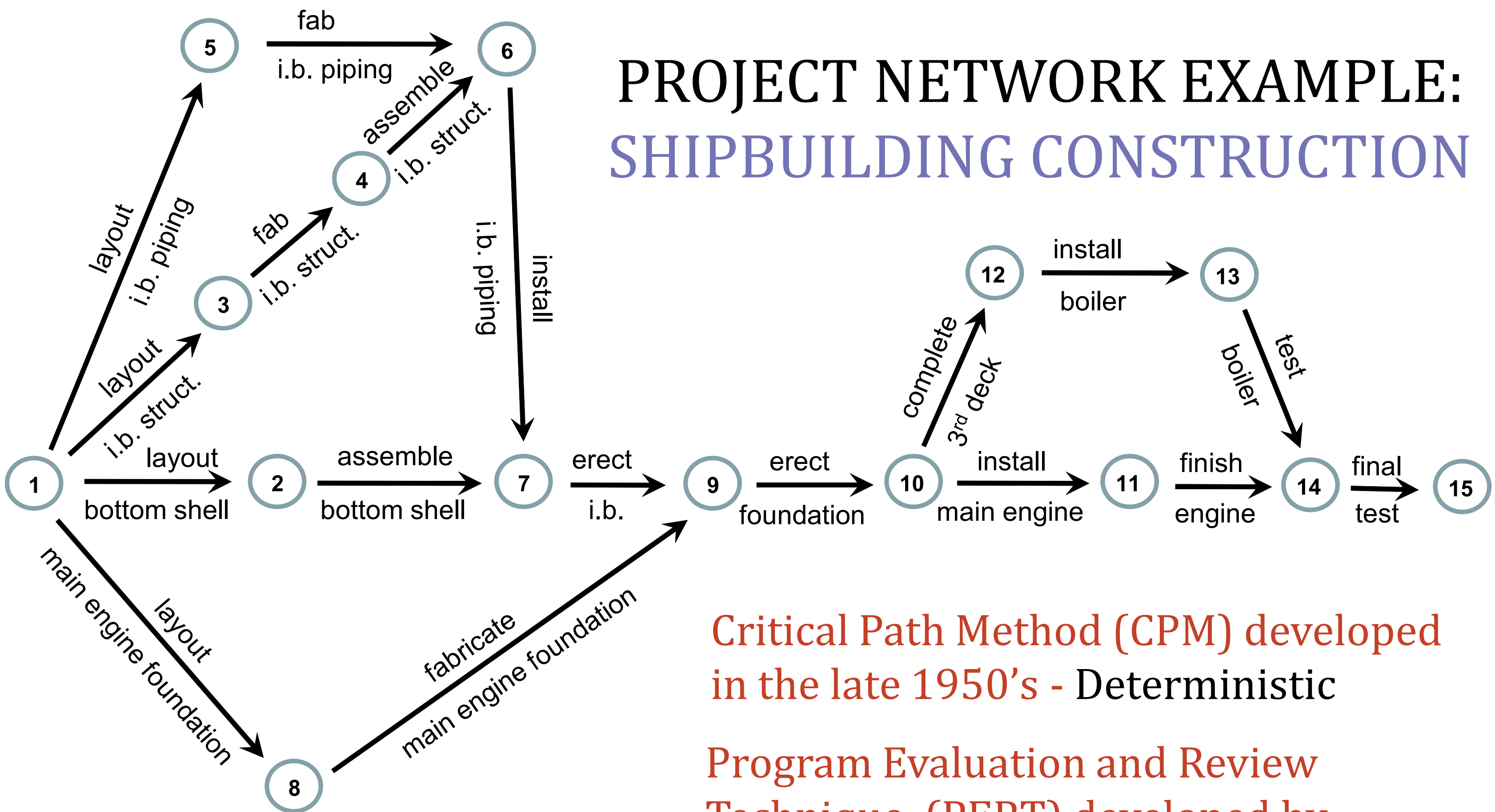
***November 20th, 2018**, La Sapienza*

J. René van Dorp*

** Department of Engineering Management and Systems Engineering,*
School of Engineering and Applied Science,The George Washington University,
800 22nd Street, N.W., Suite 2800,
Washington D.C. 20052. E-mail: dorpjr@gwu.edu
http://www.seas.gwu.edu/~dorpjr

OUTLINE

1. **INTRODUCTION**
2. ACTIVITY DURATION UNCERTAINTY MODEL
3. BAYES NETWORK STATISTICAL DEPENDENCE MODEL
4. BAYES NETWORK DEGREE OF DEPENDENCE ANALYSIS
5. DEPENDENT PERT EXAMPLE
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

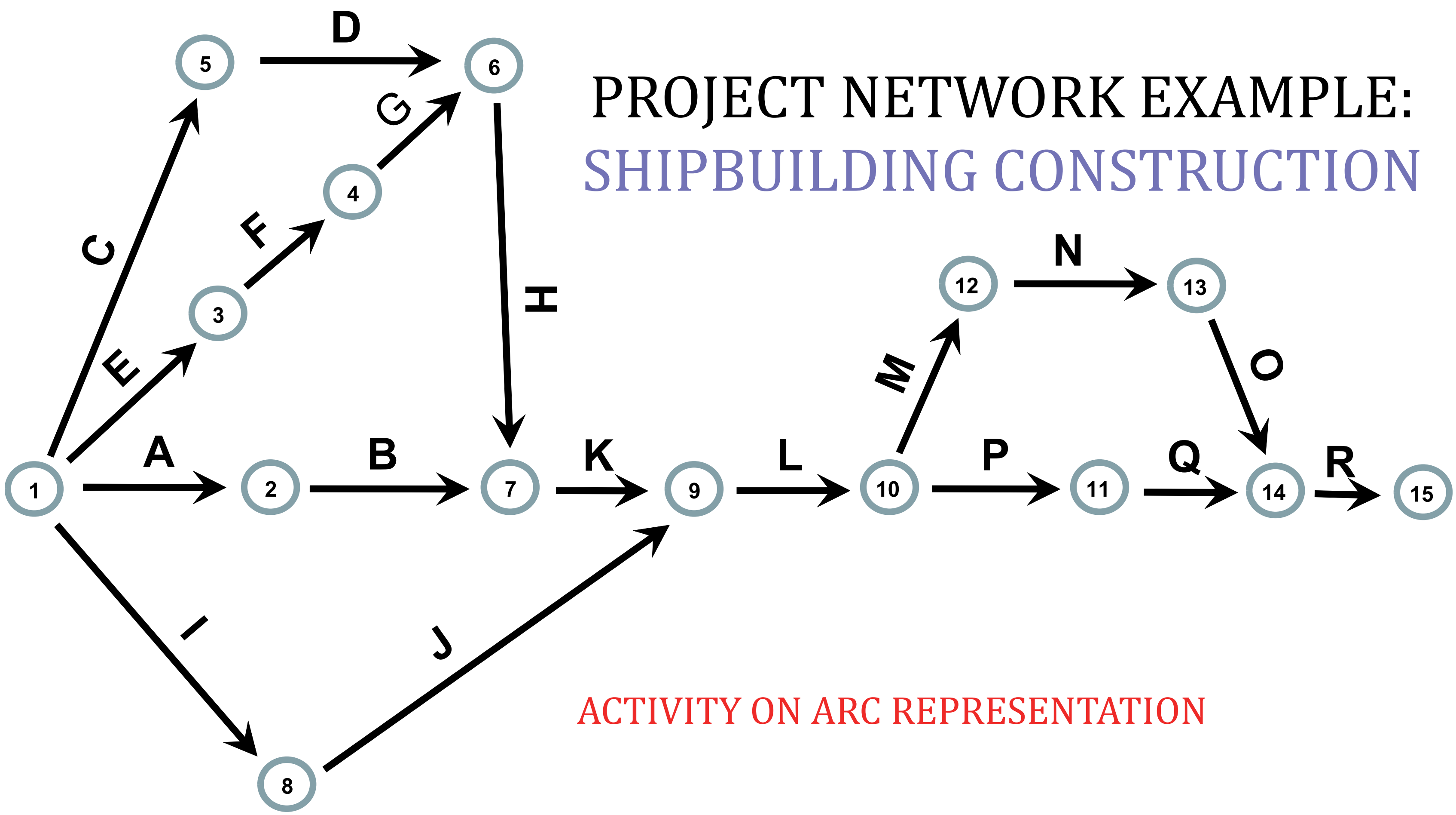


PROJECT NETWORK EXAMPLE: SHIPBUILDING CONSTRUCTION

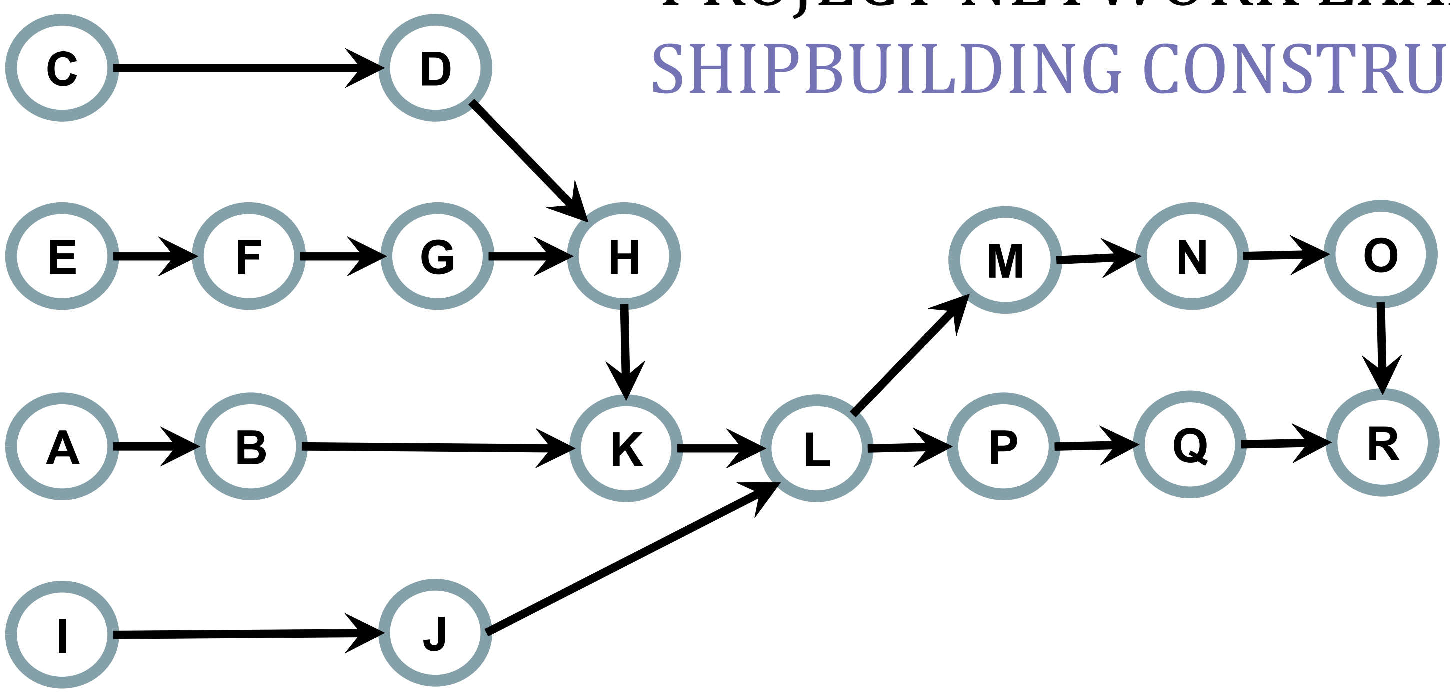
Critical Path Method (CPM) developed in the late 1950's - Deterministic

Program Evaluation and Review Technique (PERT) developed by Malcolm et al. (1959) – Stochastic CPM

PROJECT NETWORK EXAMPLE:
SHIPBUILDING CONSTRUCTION



ACTIVITY ON ARC REPRESENTATION



PROJECT NETWORK EXAMPLE: SHIPBUILDING CONSTRUCTION

ACTIVITY ON NODE REPRESENTATION

- **An activity's X duration uncertainty** in Malcolm's et al. (1959) PERT is **a beta distribution** with the following PERT mean and variance restriction

$$E[X|a_x, m_x, b_x] = \frac{a_x + 4m_x + b_x}{6}, Var[X|a_x, b_x] = \frac{(b_x - a_x)^2}{36}.$$

  Malcolm's et al. (1959) also **assumed statistical independence between the activity uncertainty durations** (albeit implicitly).

- Since Malcolm's et al. (1959) paper, **many suggestions for improvements** have been made. They center around **two main sources of criticism.**

- **The first source of criticism** relates **to using $E[X|a_x, m_x, b_x]$ and $Var[X|a_x, b_x]$ to specify the parameters** of **a beta distribution or an alternative uncertainty distribution** (see, e.g, Sasieni, 1986; Golenko-Ginzburg, 1988; Littlefield & Randolph, 1987; Gallagher, 1987; Kamburowski, 1997; Herrerias-Pleguezuelo et al., 2003; Hahn, 2008 and Herrerias-Velasco et al., 2011).

PROJECT NETWORK EXAMPLE:
SHIPBUILDING CONSTRUCTION

ID	a	m	b	δ	C(δ)	PERT Variance	Modified PERT Variance
A	22	25	30	0.375	1.250	1.778	2.222
B	35	37	43	0.250	1.143	1.778	2.032
C	19	22	29	0.300	1.194	2.778	3.317
D	4	5	10	0.167	1.032	1.000	1.032
E	23	26	31	0.375	1.250	1.778	2.222
F	16	18	24	0.250	1.143	1.778	2.032
G	11	14	20	0.333	1.222	2.250	2.750
H	6	7	12	0.167	1.032	1.000	1.032
I	25	28	33	0.375	1.250	1.778	2.222
J	33	35	40	0.286	1.181	1.361	1.607
K	27	30	37	0.300	1.194	2.778	3.317
L	6	7	11	0.200	1.080	0.694	0.750
M	4	5	9	0.200	1.080	0.694	0.750
N	6	7	10	0.250	1.143	0.444	0.508
O	9	10	15	0.167	1.032	1.000	1.032
P	6	7	12	0.167	1.032	1.000	1.032
Q	17	20	26	0.333	1.222	2.250	2.750
R	13	15	20	0.286	1.181	1.361	1.607
					Average Variance		1.790

Malcolm et al. 1959

Traditional Activity Estimates

$$\text{PERT MEAN} = \frac{a+4m+b}{6}$$

$$\text{PERT VARIANCE} = \frac{(b-a)^2}{36}$$

Herrerias et al. 2011

$$\text{MOD. PERT VARIANCE} = C(\delta)\times\frac{(b-a)^2}{36}$$

$$\delta = \frac{m-a}{b-a}, C(\delta) = \frac{5}{7}+\frac{16}{7}\times\delta(1-\delta)$$

- **The second source of criticism** relates to **the statistical independence assumption** between the activity uncertainty distributions.

- **Mathematical convenience** was **the prime motivation for the assumption of statistical independence in PERT**, despite it **not matching** the common phenomenon in practice that **multiple activities in a project network may be delayed**.

- Its mathematical convenience is derived from allowing:

  (a) for **the separate sampling from the activity marginal uncertainty distributions**, and

  (b) **evaluation of remaining project uncertainty by simply substituting observed completion times of activities in the project network structure** while **not having to modify the remaining activity uncertainty distributions** in the project network.

- Suggestions for **relaxing the statistical independence assumptions** have been made by, e.g., Jenzarli, 1994; Covaliu & Soyer, 1997; Van Dorp & Duffey, 1999; Virto et al., 2002; Cho & Covaliu, 2003; Van Dorp, 2005; Khodakarami et al., 2007; Cho, 2009; Fang & Marle, 2012.

- These **suggestions for incorporating statistical dependence** have thus far suffered from either:

  (i) **too many dependence parameters** having to be specified,
  (ii) not allowing for **the coherent monitoring of remaining project completion time uncertainty** as the project progresses,
  (iii) if (ii) is not in effect, **relying on complex numerical analysis** for that coherent monitoring.

- The purpose of this paper/presentation is **to relax that independence assumption** in **a "pragmatic" manner** that builds **on Malcolm's et al. (1959) work while attempting to address (i), (ii) and (iii) above.**

OUTLINE

1. INTRODUCTION
2. **ACTIVITY DURATION UNCERTAINTY MODEL**
3. BAYES NETWORK STATISTICAL DEPENDENCE MODEL
4. BAYES NETWORK DEGREE OF DEPENDENCE ANALYSIS
5. DEPENDENT PERT EXAMPLE
6. SUMMARY AND CONCLUSIONS
7. SELECTED REFERENCES

- **Aim:** Construct **a three-parameter uncertainty model** that has **the beta distibution and the TSP distributions as members** and has **infinitly many members that can match the PERT mean and PERT variance**.

- Duration X^* with **given lower and upper bounds a_x, b_x**, **mode M_x** $\in (a_x, b_x)$, and **shape parameter $\beta_x > 0$.** Assume:

$$X^* \sim Beta(a_x, M_x, b_x, \beta_x),$$

- **Standardizing X^*** one obtains the random variable X with support $[0, 1]$ via the linear transformation $X = (X^* - a_x)/(b_x - a_x)$ with **probability density function (pdf)** where $0 < \alpha_x < 1,\ \beta_x > 0$ and

$$g(x|\alpha_x, \beta_x) = \frac{\Gamma(\beta_x)}{\Gamma(\beta_x \alpha_x)\Gamma(\beta_x(1-\alpha_x))} \times x^{\beta_x \alpha_x - 1}(1-x)^{\beta_x(1-\alpha_x)-1},$$

- **From $g(x|\alpha_x, \beta_x)$ it immediately follows that:**

$$E[X|\alpha_x, \beta_x] = \alpha_x,\ V[X|\alpha_x, \beta_x] = \frac{\alpha_x(1-\alpha_x)}{\beta_x + 1}$$

- **Requiring $\beta_x > 2$ for $g(x|\alpha_x, \beta_x)$ to be unimodal** and **introducing the mode relative distance Δ_x of M_x** given the support $[a_x, b_x]$, one obtains :

$$\Delta_x = (M_x - a_x)/(b_x - a_x),\ 0 < \Delta_x < 1.$$

- Now **reparameterizing $g(x|\alpha_x, \beta_x)$** in terms of **its relative mode location $\Delta_x = \frac{\alpha_x \beta_x - 1}{\beta_x - 2}$** and **shape parameter $\gamma_x = \beta_x - 2 > 0$** yields

$$g(x|\Delta_x, \gamma_x) = \frac{\Gamma(\gamma_x + 2)}{\Gamma(\gamma_x \Delta_x + 1)\Gamma(\gamma_x(1 - \Delta_x) + 1)} \times x^{\gamma_x \Delta_x}(1 - x)^{\gamma_x(1-\Delta_x)}.$$

- **From $g(x|\Delta_x, \gamma_x)$ it immediately follows that:**

$$E[X|\gamma_x, \Delta_x] = \frac{\gamma_x \Delta_x + 1}{\gamma_x + 2},$$

$$V[X|\gamma_x, \Delta_x] = \frac{\gamma_x + 1 + \gamma_x^2 \Delta_x(1 - \Delta_x)}{(\gamma_x + 3)(\gamma_x + 2)^2}.$$

- **Model uncertainty in the mode relative distance such that** $\Delta_x \sim TSP(\delta_x, n_x), 0 < \delta_x < 1,\ n_x > 0,$ with pdf:

$$f(u|\delta_x, n_x) = n_x \times \begin{cases} (\frac{u}{\delta_x})^{n_x - 1}, & 0 \leq u \leq \delta_x, \\ (\frac{1-u}{1-\delta_x})^{n_x - 1}, & \delta_x \leq u \leq 1. \end{cases}$$

- **The cumulative distribution function (cdf)** for the pdf above is

$$F(u|\delta_x, n_x) = \begin{cases} \delta_x(\frac{u}{\delta_x})^{n_x}, & 0 \leq u \leq \delta_x, \\ 1 - (1-\delta_x)(\frac{1-u}{1-\delta_x})^{n_x}, & \delta_x \leq u \leq 1, \end{cases}$$

- **The quantile function** for the pdf above is

$$F^{-1}(y|\delta_x, n_x) = \begin{cases} \delta_x(\frac{y}{\delta_x})^{1/n_x}, & 0 \leq y \leq \delta_x, \\ 1 - (1-\delta_x)(\frac{1-y}{1-\delta_x})^{1/n_x}, & \delta_x \leq y \leq 1. \end{cases}$$

- Observe from **the cdf and the qf** that:

$$F^{-1}(y|\delta_x, n_x) = F(y|\delta_x, 1/n_x).$$

- **From $f(u|\delta_x, n_x)$ it immediately follows that:**

$$E[\Delta_x|\delta_x, n_x] = \frac{(n_x - 1)\delta_x + 1}{n_x + 1},$$
$$V[\Delta_x|\delta_x, n_x] = \frac{n_x - 2(n_x - 1)\delta_x(1 - \delta_x)}{(n_x + 2)(n_x + 1)^2}.$$

- **Given that a quantile level y in the quantile function completely determines the value for Δ_x** one may write:

$$E[\Delta_x|y, \delta_x, n_x] = F(y|\delta_x, 1/n_x).$$

- **Given $g(x|\Delta_x, \gamma_x)$ for $(X|\Delta_x)$** and **the prior pdf $f(u|\delta_x, n_x)$ for Δ_x**, closed form expression for **the predictive mean** and **predictive variance for the activity duration X** can be derived by applying repeatedly **the law of total expectation $(LOTE)$ $E_X[X] = E_Y\left[E_{X|Y}[X|Y]\right]$**

- With $\Delta_x = F^{-1}(Y|\delta_x, n_x)$, where $Y \sim U[0,1]$ and applying the $LOTE$ it follows with $F^{-1}(y|\delta_x, n_x) = F(y|\delta_x, 1/n_x)$ that:

$$E[X|\gamma_x, \delta_x, n_x] = E_Y\left[E_{X|Y}[X|\gamma_x,\ \Delta_x = F^{-1}(Y|\delta_x, n_x)]\right]$$
$$= \int_{y=0}^{1} E[X|\gamma_x, F^{-1}(y|\delta_x, n_x)]dy = \int_{y=0}^{1} E[X|\gamma_x, F(y|\delta_x, 1/n_x)]dy.$$

- Substitution of $E[X|\gamma_x, \Delta_x] = (\gamma_x \Delta_x + 1)/(\gamma_x + 2)$ for X yields

$$E[X|\gamma_x, \delta_x, n_x] = \frac{\gamma_x \int_{y=0}^{1} F(y|\delta_x, 1/n_x)dy + 1}{\gamma_x + 2}.$$

- Utilizing $Y \sim U[0,1]$, $F(y|\delta_x, 1/n_x) = E[\Delta_x|y, \delta_x, n_x]$ yields

$$E[X|\gamma_x, \delta_x, n_x] = \frac{\gamma_x \int_{y=0}^{1} E[\Delta_x|y, \delta_x, n_x]dy + 1}{\gamma_x + 2}$$
$$= \frac{\gamma_x E[\Delta_x|\delta_x, n_x] + 1}{\gamma_x + 2} = E\big[X|\gamma_x, E[\Delta_x|\delta_x, n_x]\big].$$

- Analogously, **using similar algebraic manipulations** it follows

$$Var[X|\gamma_x,\delta_x,n_x] = Var\big[X|\gamma_x, E[\Delta_x|\delta_x,n_x]\big] + \frac{\gamma_x^2}{(\gamma_x+2)(\gamma_x+3)} \times Var[\Delta_x|\delta_x,n_x].$$

- Thus, the predictive variance $Var[X|\gamma_x,\delta_x,n_x]$ equals **the sum** of **the variance $Var\big[X|\gamma_x, E[\Delta_x|\delta_x,n_x]\big]$**, where **the prior modal mean $E[\Delta_x|\delta_x,n_x]$ plays the role of the parameter Δ_x** , **plus** the **prior modal variance $Var[\Delta_x|\delta_x,n_x]$**, curtailed by $0 < \gamma_x^2/(\gamma_x+2)(\gamma_x+3) < 1$.

- Similarly, **predictive pdf can be approximated** using the fact that $\Delta_x = F^{-1}(Y|\delta_x,n_x)$, where $Y \sim U[0,1]$ and $F^{-1}(y|\delta_x,n_x) = F(y|\delta_x,1/n_x)$ by

$$h(x|\delta_x,n_x,\gamma_x) = \int_{u=0}^{1} g(x|u,\gamma_x)f(u|\delta_x,n_x)du \approx \frac{1}{(n+1)}\sum_{i=0}^{n} g(x|F(i/n|\delta_x,1/n_x),\gamma_x).$$

- **Summarizing,** a three parameter uncertainty model $h(x|\delta_x, n_x, \gamma_x)$ for the activity duration uncertainty X has been constructed **combining a beta distribution** and **a TSP distribution**.

- **Given a specified value for γ_x**, values for n_x and δ_x can be solved such that:

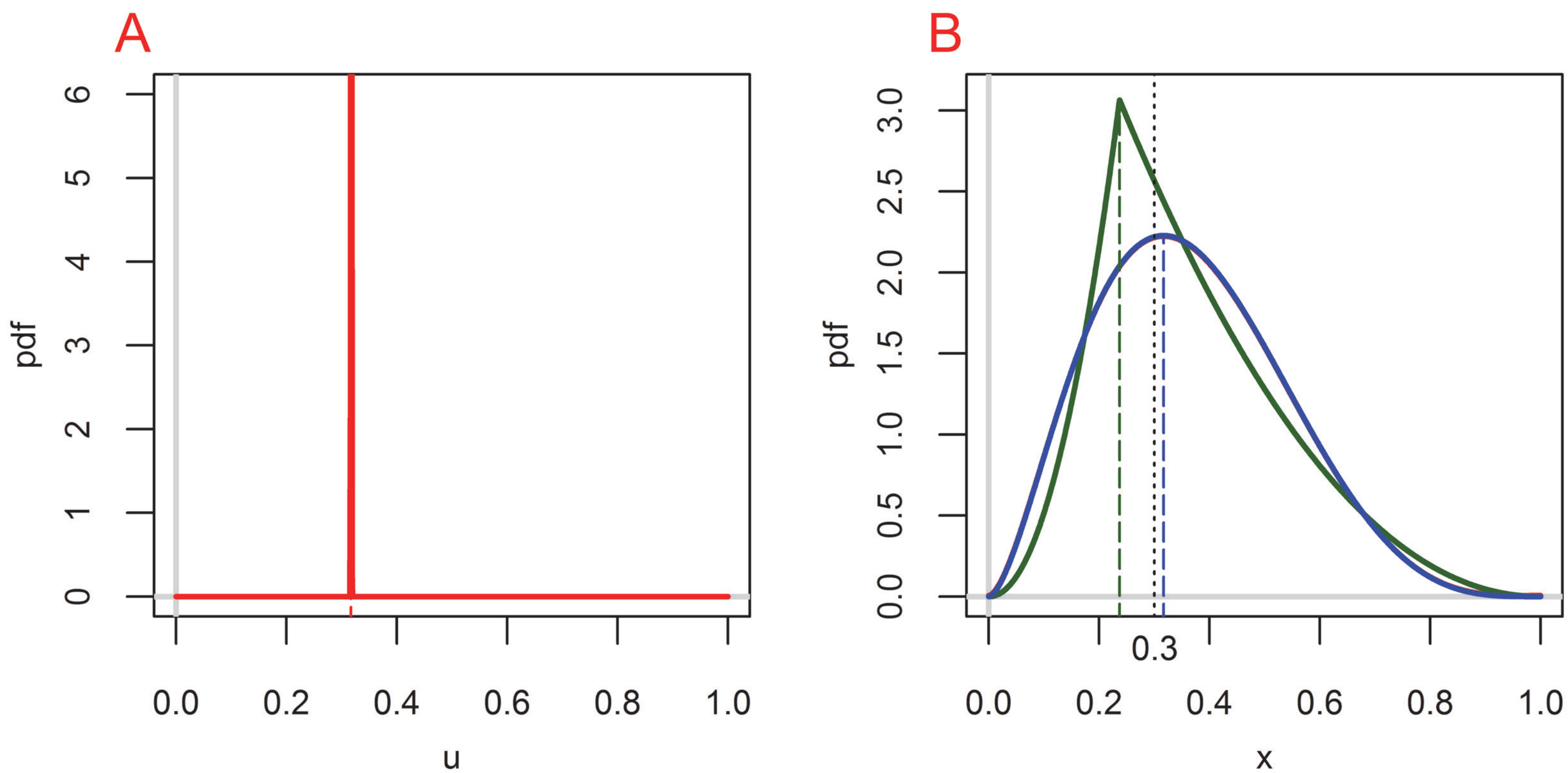
$$E[X|\delta_x, n_x, \gamma_x] = \frac{0 + 4 \times m_x + 1}{6}, V[X|\delta_x, n_x, \gamma_x] = \frac{(1-0)^2}{36}$$

- **For this given value of γ_x and solved values for n_x and δ_x**, and using the linear transformation $X^* = (b_x - a_x)X + a_x$ it follows that:

$$E[X|\delta_x, n_x, \gamma_x] = \frac{a_x + 4 \times m_x^* + b_x}{6}, V[X|\delta_x, n_x, \gamma_x] = \frac{(b_x - a_x)^2}{36}$$

- **Conclusion:** an uncertainty duration model has been constructed where **infinitly many of its members** can match the PERT mean and PERT variance. Recall, only **one single beta member distribution** and only **one single TSP member distribution** can match the PERT mean and PERT variance.

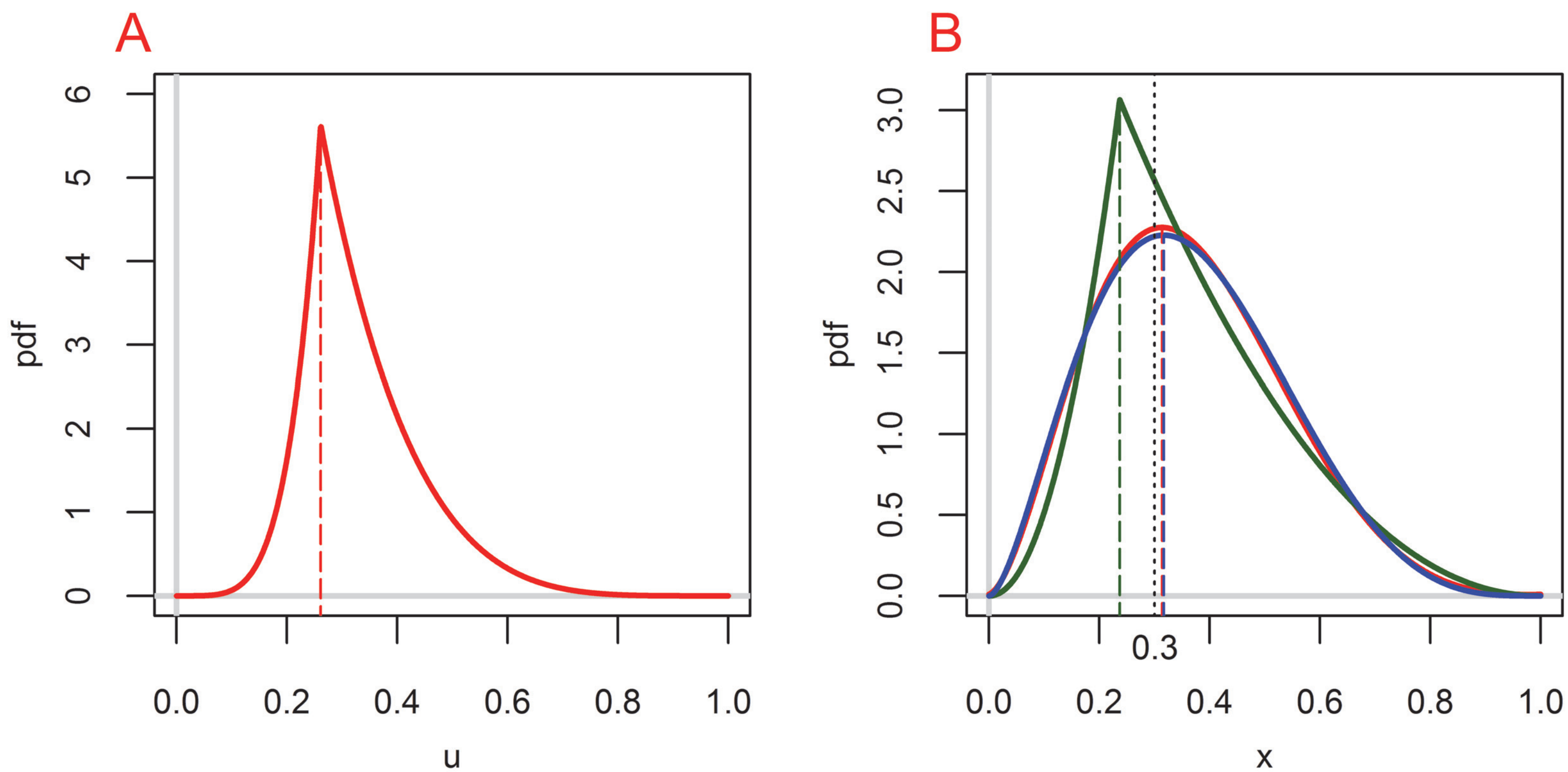
$f(u|\delta_x, n_x)$ $h(x|\delta_x, n_x, \gamma_x)$



$\delta_x = 0.3168, n_x \to \infty$ $\gamma_x = 5.36$

- **All "three" probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

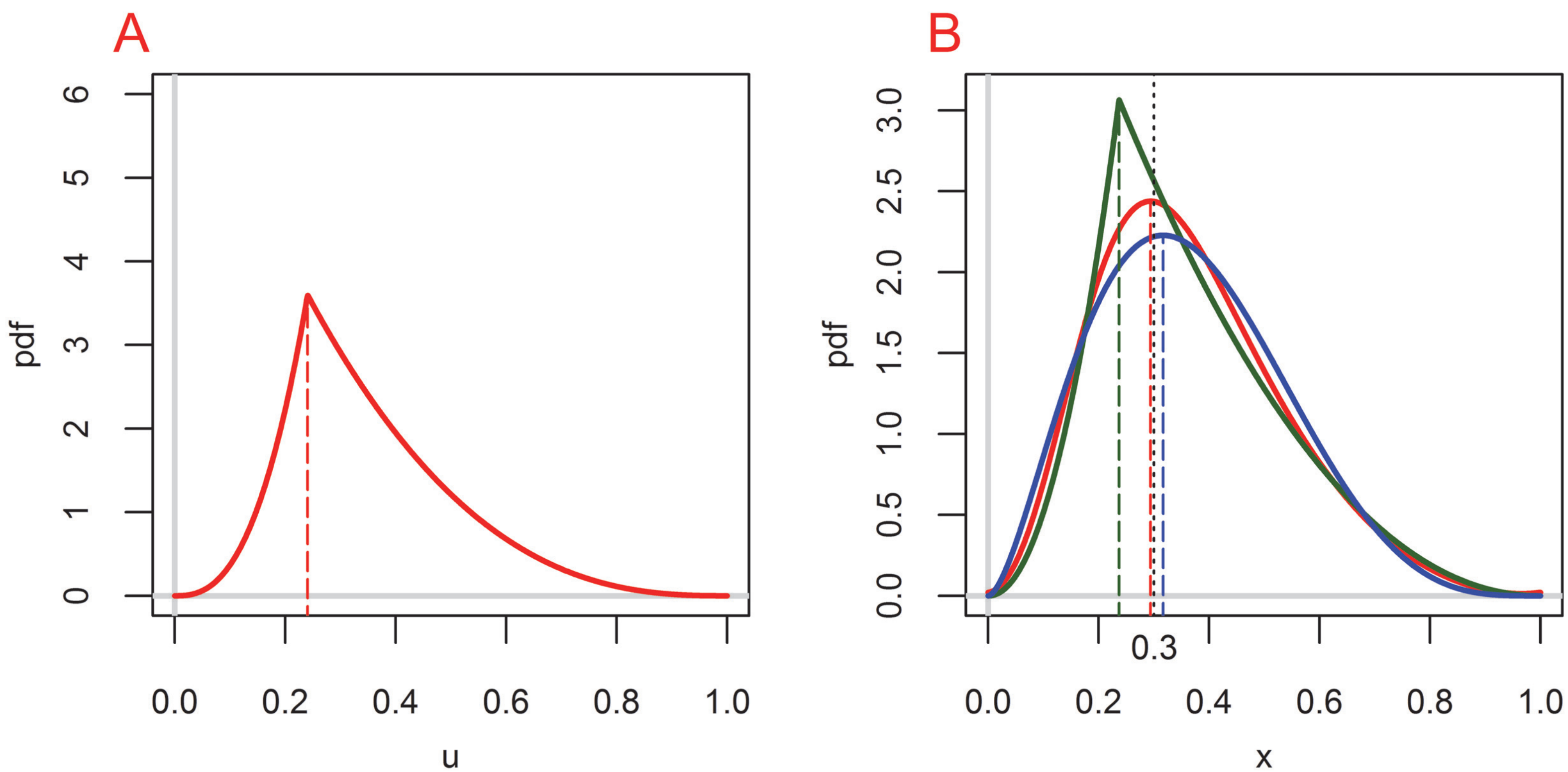
$f(u|\delta_x, n_x)$ $h(x|\delta_x, n_x, \gamma_x)$



$\delta_x = 0.2613, n_x = 5.6295$ $\gamma_x = 8$

- **All three probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

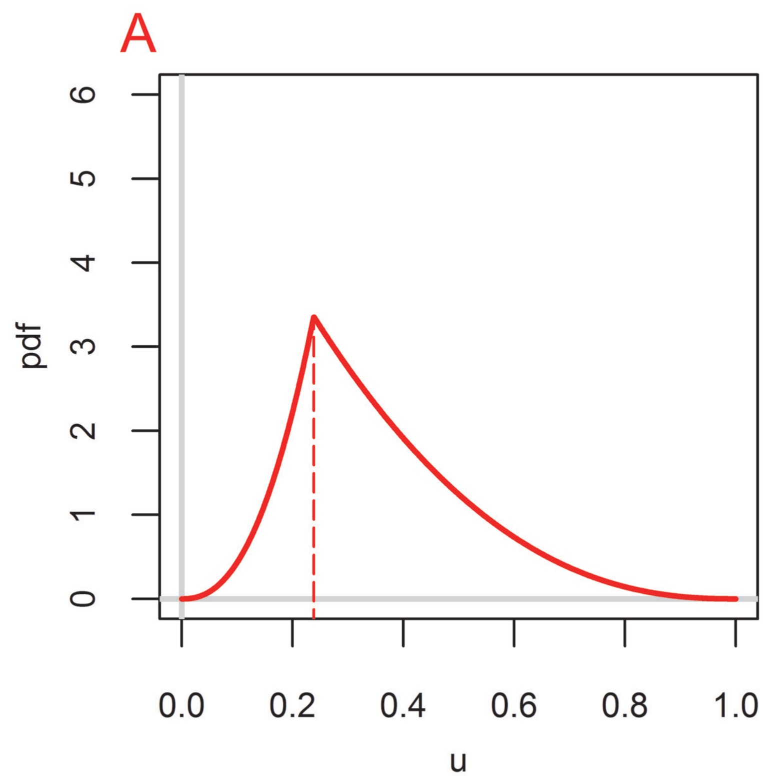
$f(u|\delta_x, n_x)$ $h(x|\delta_x, n_x, \gamma_x)$



$\delta_x = 0.2426, n_x = 3.7939$ $\gamma_x = 16$

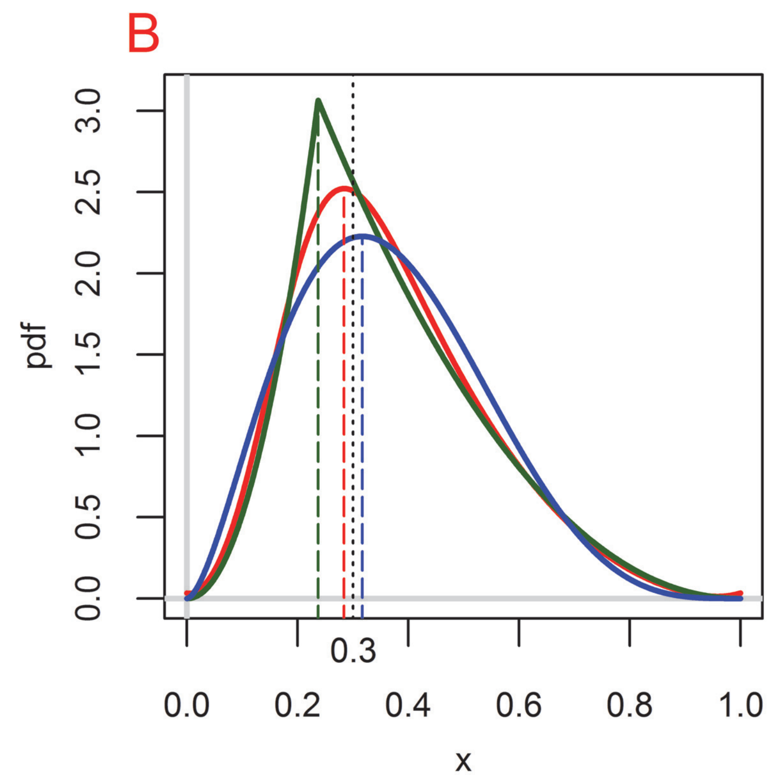
- **All three probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

$f(u|\delta_x, n_x)$



$\delta_x = 0.2382, n_x = 3.3598$

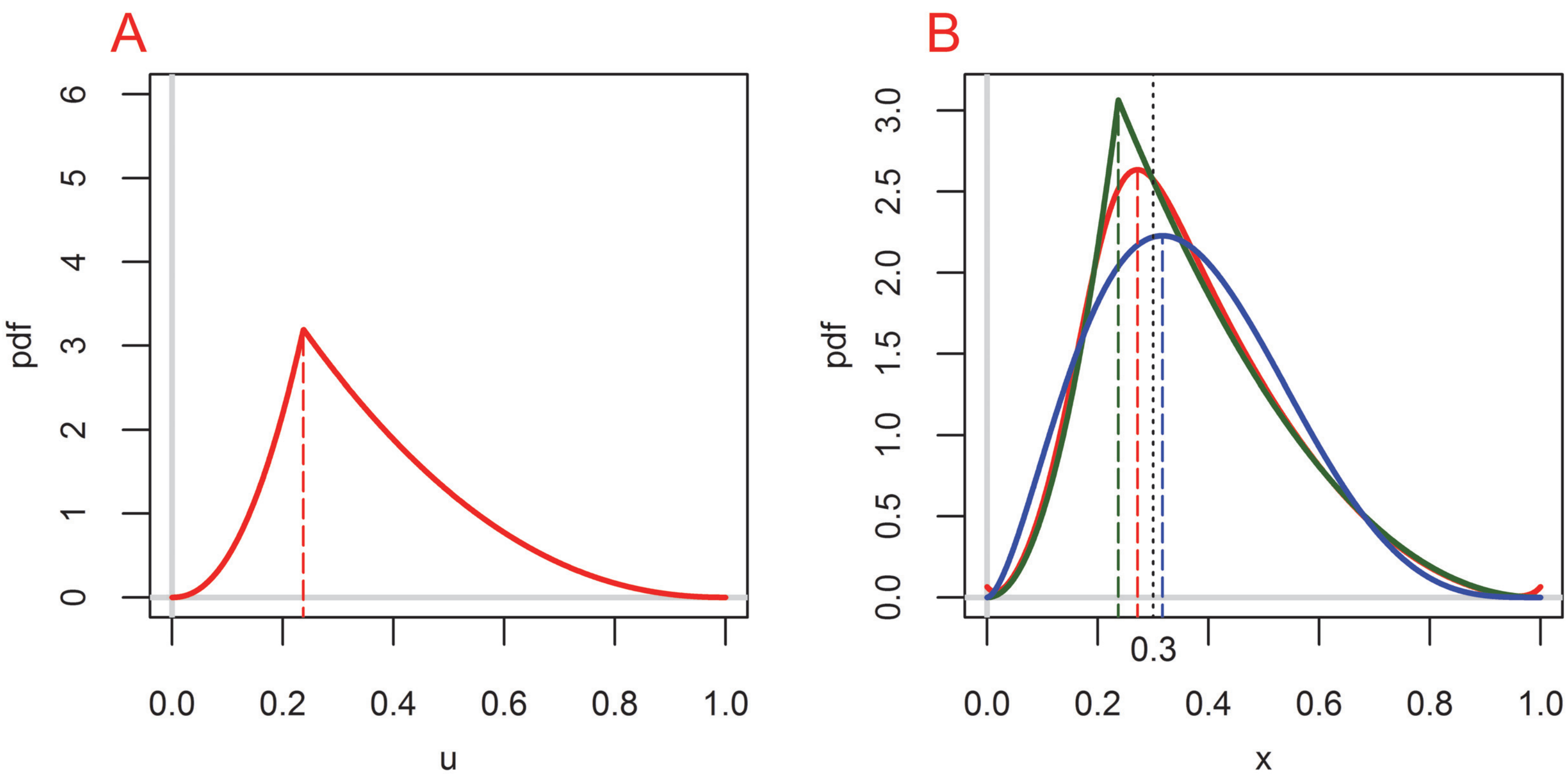
$h(x|\delta_x, n_x, \gamma_x)$



$\gamma_x = 32$

- **All three probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

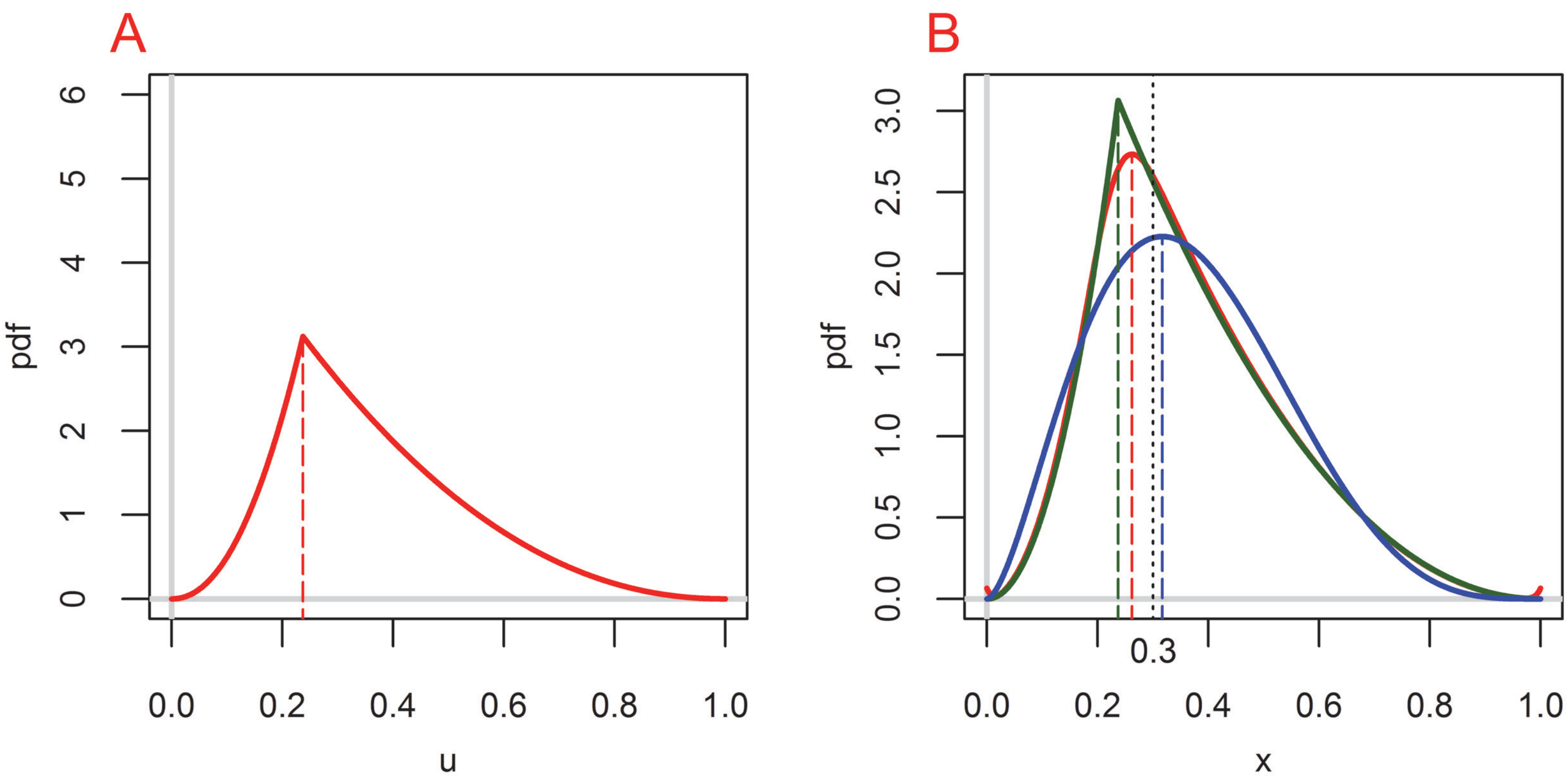
$f(u|\delta_x, n_x)$ $h(x|\delta_x, n_x, \gamma_x)$



$\delta_x = 0.2373, n_x = 3.1960$ $\gamma_x = 64$

- **All three probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

$f(u|\delta_x, n_x)$ $h(x|\delta_x, n_x, \gamma_x)$

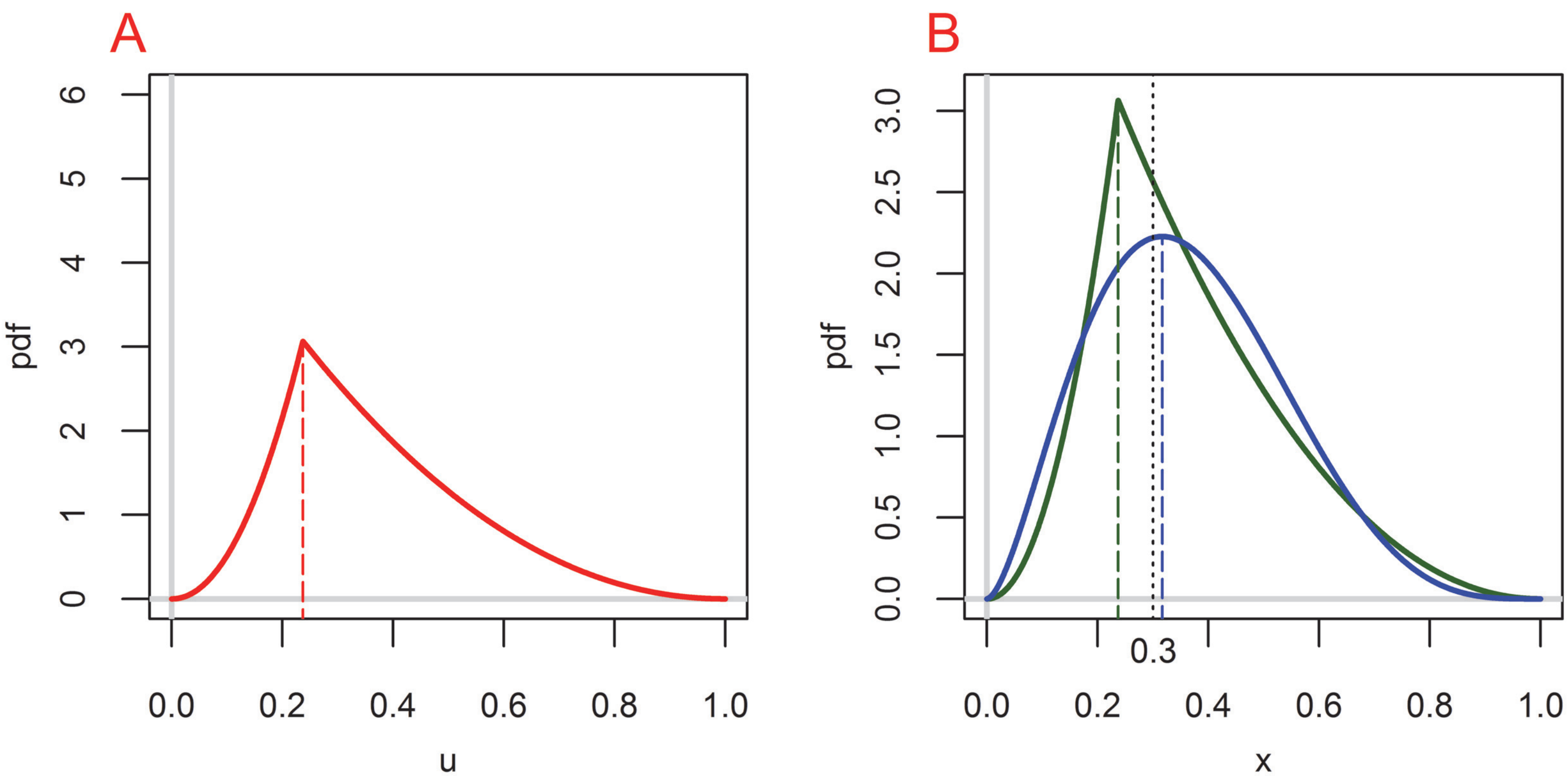


$\delta_x = 0.2371, n_x = 3.1240$ $\gamma_x = 128$

- **All three probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

2. ACTIVITY DURATION... Predictive Distribution of X

$f(u|\delta_x, n_x)$ $\qquad$ $h(x|\delta_x, n_x, \gamma_x)$



$\delta_x = 0.2371, n_x = 3.0619$ $\qquad$ $\gamma_x \rightarrow \infty$

- **All three probability density functions** in Figure B **match the PERT mean and PERT variance**. Pdf in **blue** (**green**) in Figure B is the **beta PERT** (**TSP PERT**) distribution with $a_x = 0, m_x = 0.3, b_x = 1$.

OUTLINE

1. INTRODUCTION

2. ACTIVITY DURATION UNCERTAINTY MODEL

3. **BAYES NETWORK STATISTICAL DEPENDENCE MODEL**

4. BAYES NETWORK DEGREE OF DEPENDENCE ANALYSIS

5. DEPENDENT PERT EXAMPLE

6. SUMMARY AND CONCLUSIONS

7. SELECTED REFERENCES

- Let A^*, B^* be **two activities X^* as described previously**, in a project network, and let A and B be their standardized versions with support $[0, 1]$.

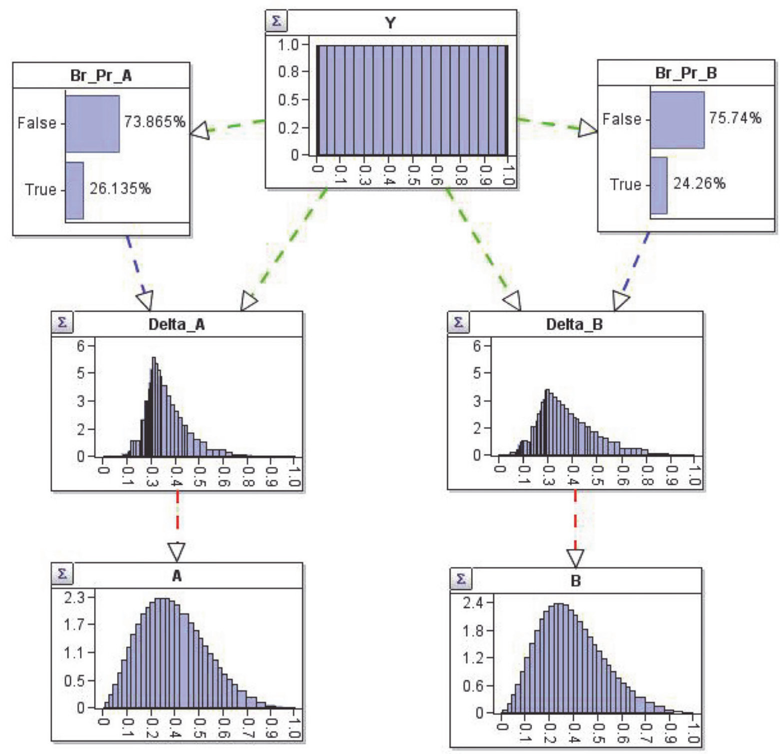
- **The suggested Bayesian Network (BN) model** to the right is a screen shot from **the Bayes Network software AgenaRisk®.**

Common Mode Quantile Level for A and B:

$$Y \sim U[0,1]$$

$$(\Delta_a|Y) = F^{-1}(Y|\delta_a, n_a),$$
$$(\Delta_b|Y) = F^{-1}(Y|\delta_b, n_b).$$

$$(A|\Delta_a) \sim g(\,\cdot\,|\Delta_a, \gamma_a),$$
$$(B|\Delta_b) \sim g(\,\cdot\,|\Delta_b, \gamma_b).$$



- This BN model **naturally extends to more than two activities**.

- Observe from figure that **the random variables $\Delta_a \sim TSP(\delta_a, n_a)$ and $\Delta_b \sim TSP(\delta_b, n_b)$** are **conditionally independent given $Y \sim U[0,1]$**.

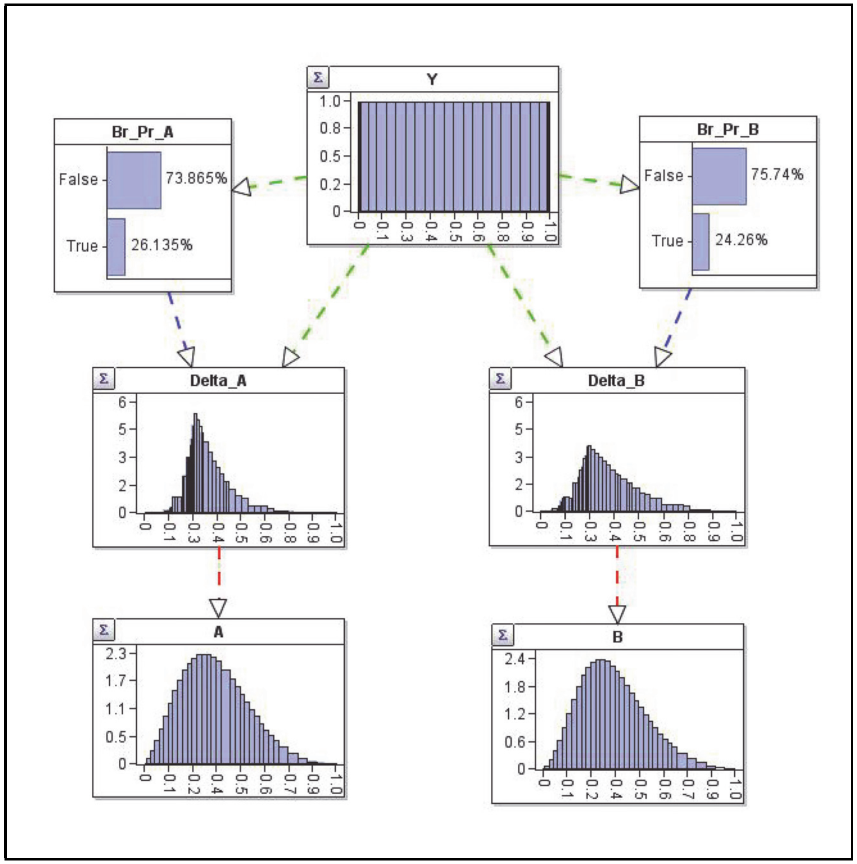
- Observe from figure above that **the random variables A and B** are **conditionally independent given Y** or, alternatively, **the random variables A and B** are **conditionally independent given $(\Delta_a,\ \Delta_b)$**.

- **The node $Y \sim U[0,1]$ is the common quantile level node** for **the mode relative distances $(\Delta_a,\ \Delta_b)$.** Thus **the modes of the activity durations A and B "*move-in-sync*"** as per their common quantile level Y.

- Thus given a **high** (**low**) common quantile level, **both uncertainty distributions for A and B will be skewed** towards the **right** (**left**). As a result activity durations **A and B will be positively dependent**.

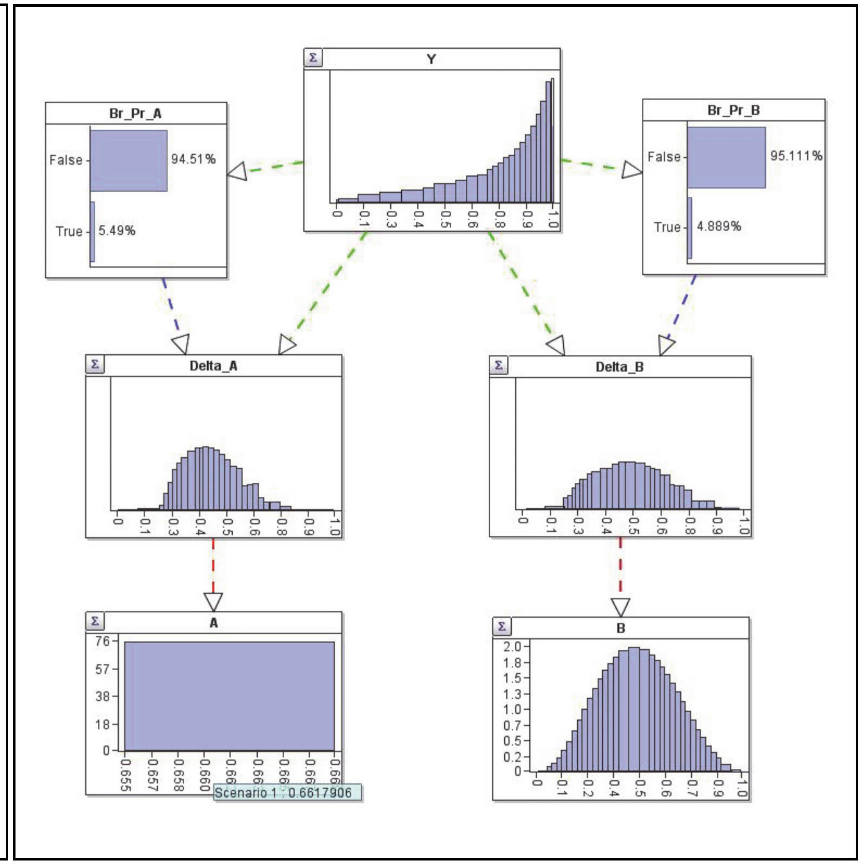
- **Their degree of dependence as measured by the correlation** between A and B will be a function of **the relative mode location parameters δ_x, n_x** and **the activity duration parameters γ_x**, where $x \in \{a, b\}$.

Prior Distribution



No Instantiation of A or B

Posterior Distribution



Activity A instantiated with $a_{0.95}$

- **Posterior updating** performed by software **AgenaRisk®.** Note **the shift towards the right** for B, Δ_a, Δ_b and Y. Also, $\boldsymbol{E[Y|a_{0.95}] \approx 0.95}$.

OUTLINE

- Utilizing **conditional independence of Δ_a and Δ_b given $Y \sim U[0,1]$** and **the law of total expectation** it can be shown:

$$
\begin{aligned}
Cov[\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b] &= \frac{1+(n_b-1)\delta_b}{(n_b+1)}\left[1 - \frac{1+(n_a-1)\delta_a}{n_a+1}\right] \\
&+ \left[\frac{n_a n_b}{n_a+n_b+n_a n_b} - \frac{n_b}{n_b+1}\frac{1}{\delta_a}\right](\delta_a)^2 \delta_b \sqrt[n_b]{\frac{\delta_a}{\delta_b}} \\
&+ \left[\frac{n_a n_b}{n_a+n_b+n_a n_b} - \frac{n_a}{n_a+1}\frac{1}{1-\delta_b}\right](1-\delta_a)(1-\delta_b)^2 \times \sqrt[n_a]{\frac{1-\delta_b}{1-\delta_a}} \\
&+ \frac{(1-\delta_a)\delta_b}{\sqrt[n_a]{1-\delta_a} \times \sqrt[n_b]{\delta_b}} \\
&\times \frac{[B(1-\delta_b|1/n_a+1, 1/n_b+1) - B(1-\delta_a|1/n_a+1, 1/n_b+1)]}{\Gamma(1/n_a+1/n_b+2)/\Gamma(1/n_a+1)\Gamma(1/n_b+1)}
\end{aligned}
$$

- **$Cov[\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b] \downarrow 0$ as $n_a \to \infty, n_a \to \infty$.** This also follows from **BN Structure** and the observation that when **$n_a \to \infty$, $n_a \to \infty$ the TPS pdfs for Δ_a and Δ_a converge to point masses at δ_a, δ_b.**

- Utilizing **conditional independence of A and B given $Y \sim U[0,1]$** and **the law of total expectation** it can also be shown that:

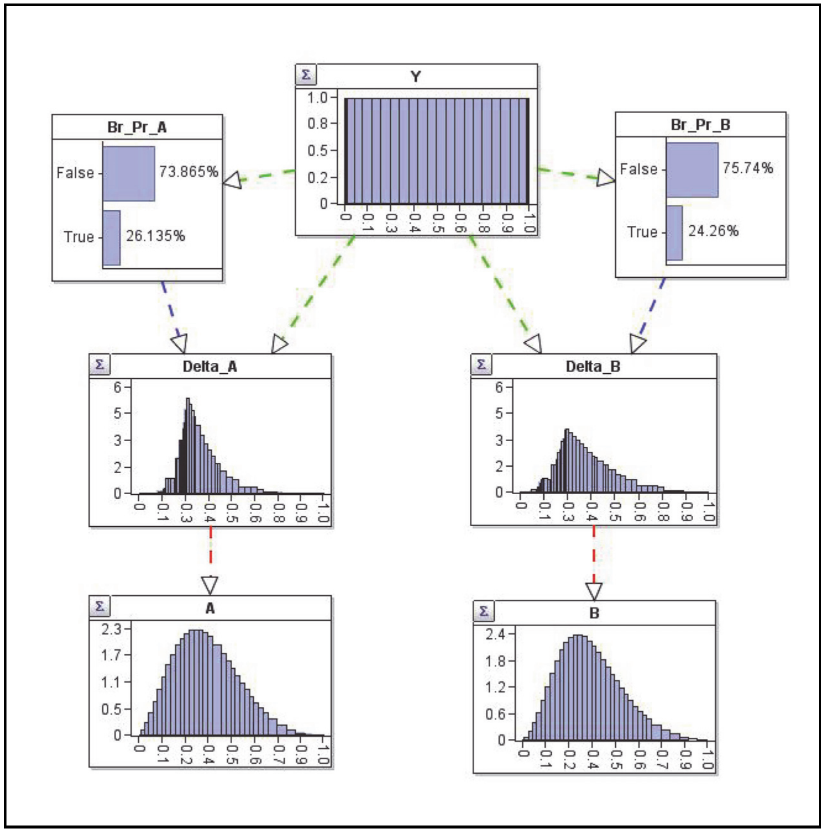
$$Cov(A,B|\Theta) = \frac{\gamma_a \gamma_b}{(\gamma_a+2)(\gamma_b+2)} \times Cov(\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b),$$

- Under the condition that **activity A and activity B adhere to the PERT mean and PERT variance**, one has $Var[A] = Var[B] = 1/36$ which yields with

$$\rho(A,B|\Theta) = \frac{36\gamma_a \gamma_b}{(\gamma_a+2)(\gamma_b+2)} \times Cov(\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b).$$

- $\gamma_a = 8 \Rightarrow n_a = 5.629,\ \delta_a = 0.261$ **for A** to match PERT mean and PERT Variance given $m_a = 0.3$.

- $\gamma_b = 16 \Rightarrow n_b = 3.794,\ \delta_b = 0.243$ **for B** to match PERT mean and PERT Variance given $m_b = 0.3$.



Step 1:

$$\begin{cases} n_a = 5.629,\ \delta_a = 0.261, & \text{for } \Delta_a \\ n_b = 3.794,\ \delta_b = 0.243, & \text{for } \Delta_b \end{cases}$$

$$\Downarrow$$

$$Cov[\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b] \approx 0.0154.$$

Step 2:

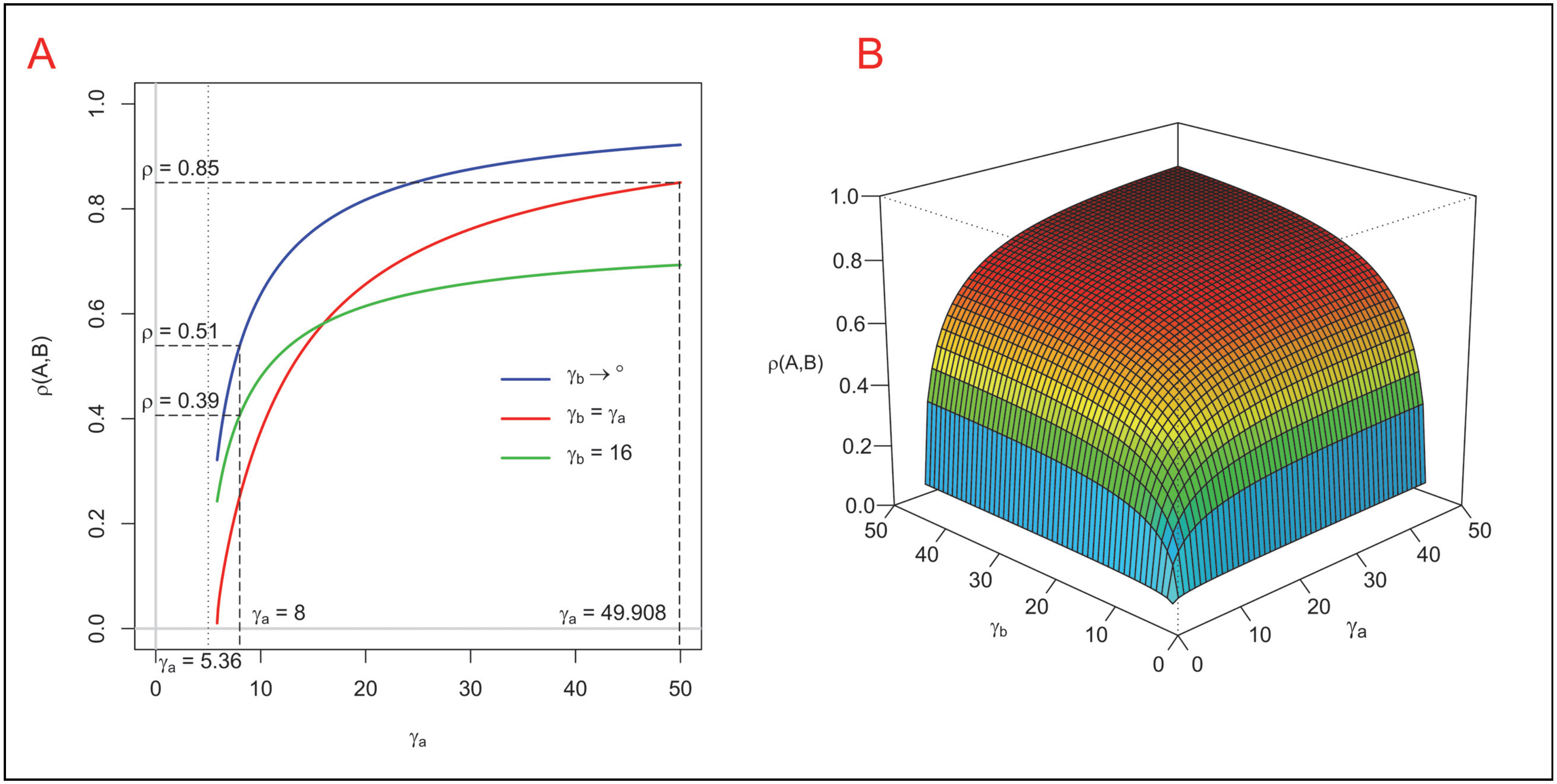
$$V[\Delta_a|\delta_a, n_a] \approx 0.0115,$$
$$V[\Delta_b|\delta_b, n_b] \approx 0.0208,$$
$$Cov[\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b] \approx 0.0154$$
$$\Downarrow$$
$$\rho[\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b] \approx 0.9986.$$

Step 3:

$$Cov[\Delta_a, \Delta_b | \delta_a, \delta_b, n_a, n_b] \approx 0.0154,$$
$$\gamma_a = 8 \text{ for activity } A$$
$$\gamma_b = 16 \text{ for activity } B$$
$$\Downarrow$$
$$Cov(A, B|\Theta) \approx 0.0110.$$

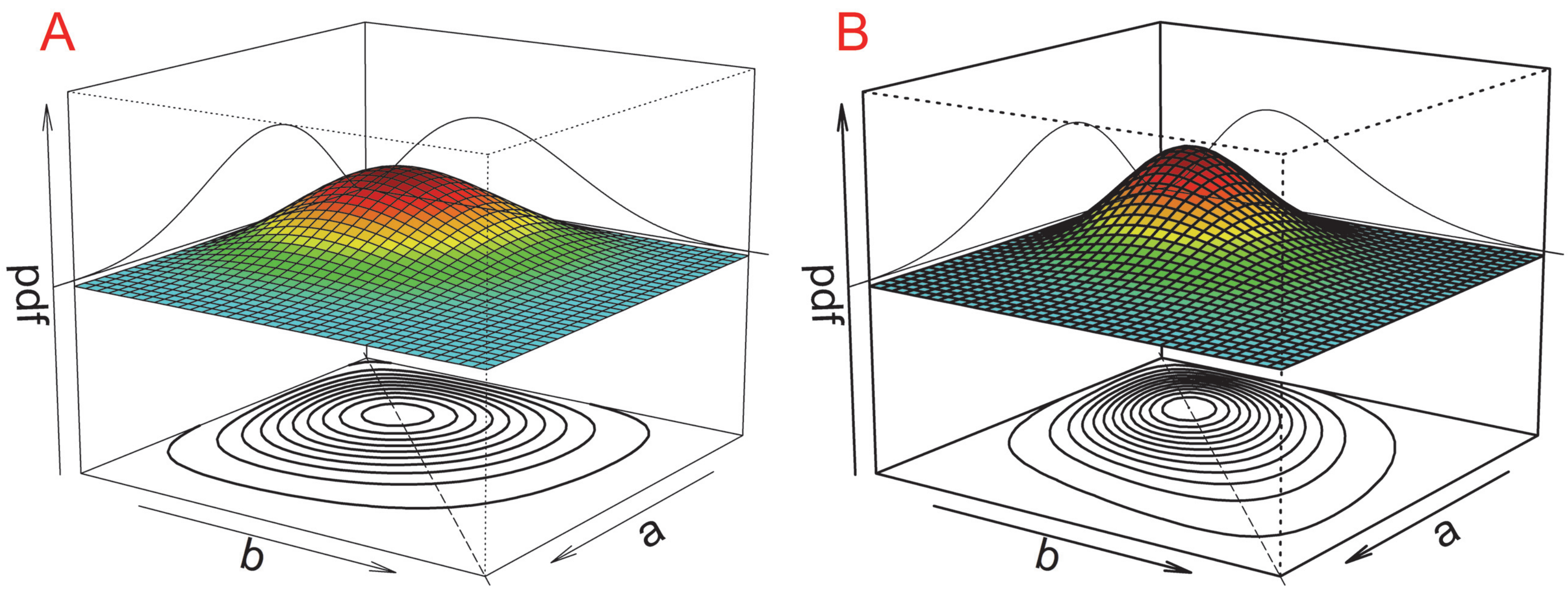
Step 4:

$$Cov(A, B|\Theta) \approx 0.0110.$$
$$Var(A) = 1/36,\ Var(B) = 1/36$$
$$\Downarrow$$
$$\rho(A, B|\Theta) \approx 0.3945.$$



Behavior of $\rho(A, B|\Theta)$ for the random variables A and B in **the Bayes Network** as a function of **the dependence parameters γ_a and γ_b** with the random variables A and B possessing the PERT mean and PERT variance given supports $[0, 1]$ and most likely estimates $m_a = m_b = 0.3$.

Predictive joint pdf $h(a, b|\Theta)$ can be approximated as the univariate ones.

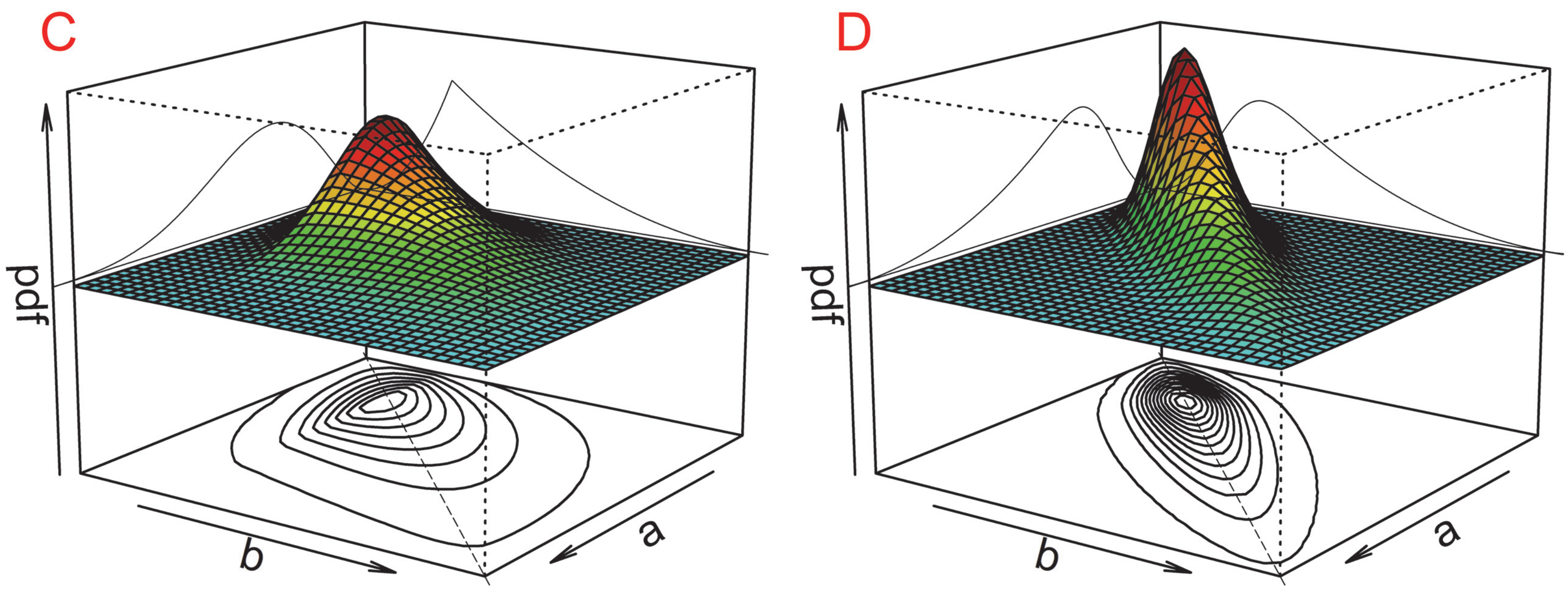


Examples of the joint pdf with most likely estimates $m_a = m_b = 0.3$ with marginal pdfs for A and B possessing PERT mean and PERT variance but different correlations:

A: $\rho(A, B|\Theta) \approx 0$, $\gamma_a = 5.36,\ n_a \to \infty,\ \delta_a = 0.317, \gamma_b = 5.36,\ n_b \to \infty,\ \delta_b = 0.317;$

B: $\rho(A, B|\Theta) \approx 0.39$, $\gamma_a = 8,\ n_a = 5.629,\ \delta_a = 0.261, \gamma_b = 16,\ n_b = 3.794,\ \delta_b = 0.243;$

Predictive joint pdf $h(a, b|\Theta)$ can be approximated as the univariate ones.



Examples of the joint pdf with most likely estimates $m_a = m_b = 0.3$ with marginal pdfs for A and B possessing PERT mean and PERT variance but different correlations:

C: $\rho(A, B|\Theta) \approx 0.51$, $\gamma_a = 8,\ n_a = 5.629,\ \delta_a = 0.261, \gamma_b \to \infty,\ n_b = 3.059,\ \delta_b = 0.237;$

D: $\rho(A, B|\Theta) \approx 0.85$, $\gamma_a = \gamma_b = 49.908,\ n_a = n_b = 3.239,\ \delta_a = \delta_b = 0.237.$

- **With A and B adhering to the PERT Mean and PERT Variance,** it follows that when $n_a \to \infty$, $n_b \to \infty$ the **joint predictive pdf** is one of **independent beta PERT distributions** with $\rho(A, B) = 0$.

- **With A and B adhering to the PERT Mean and PERT Variance,** it follows that when $\gamma_a \to \infty$, $\gamma_b \to \infty$ the **joint predictive pdf** is one of **dependent TSP PERT distributions** with $\rho(A, B) = 1$.

- **With A and B adhering to the PERT Mean and PERT Variance,** a **joint predictive pdf** can be solved for with $0 < \rho(A, B) < 1$.

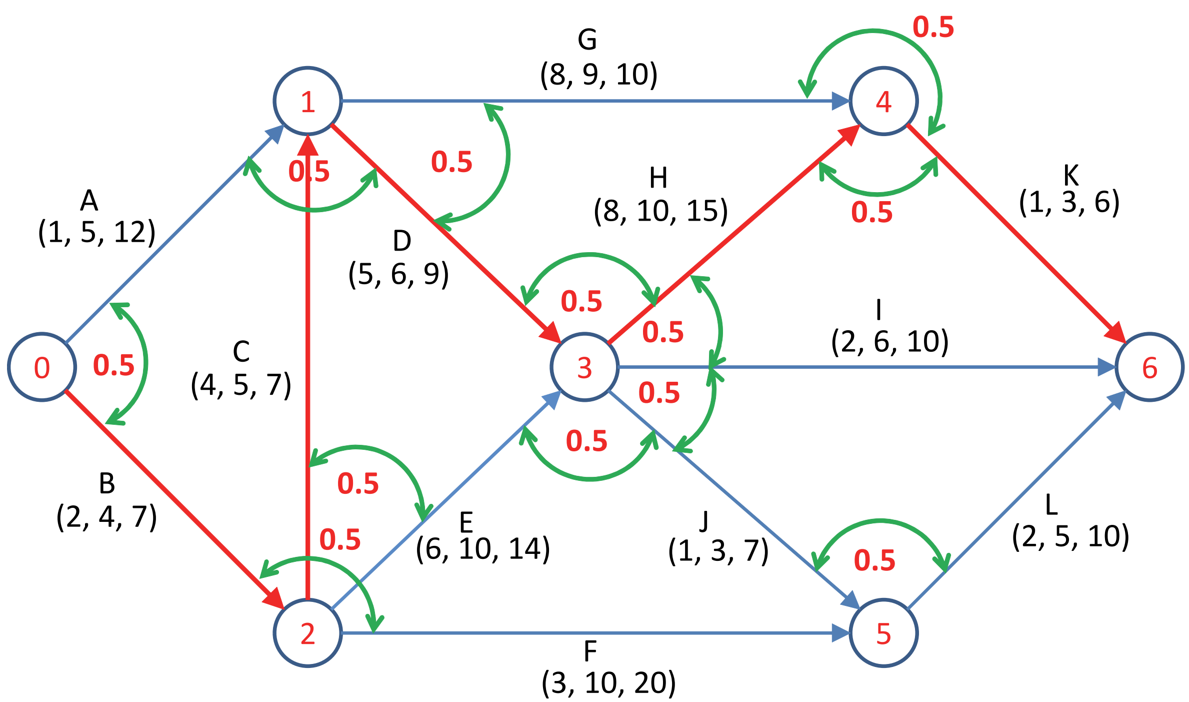
- **This joint statistical dependence model** does **not allow** for **the modeling of negative dependence**.

- For given values of γ_a, γ_b, **the PERT mean and PERT Variance restriction** determine the parameter values for n_a, n_b **and** δ_a, δ_b.

- **Conclusion:** γ_a, γ_b are **dependence parameters** of **the Bayes Network.**

OUTLINE

1. INTRODUCTION
2. ACTIVITY DURATION UNCERTAINTY MODEL
3. BAYES NETWORK STATISTICAL DEPENDENCE MODEL
4. BAYES NETWORK DEGREE OF DEPENDENCE ANALYSIS
5. **DEPENDENT PERT EXAMPLE**
6. SUMMARY AND CONCLUSUONS
7. SELECTED REFERENCES



Example 6-node project network consisting of 12 activities.

- When **using the most likely values m_x** for the activities, **the critical path has a length of 28** and is indicated using **red arrows.**

- **The same critical path** emerges when using **the lower bounds a_x** (**the lower bounds b_x**) with **a length 20** (**length 44**).

To specify parameters $(\gamma_i, \delta_i, n_i)$ for $i = 1, \ldots, n$ activities X_i in the project above, it is proposed to **select a set of n correlations ρ_{i_k, j_k} , $j_k \neq i_k$** among paired activity durations to be set equal to a **"project-degree of statistical dependence" $\rho = 0.5$** and solve :

$$\min_{\mathcal{S}} : \sum_{k=1}^{n} (\rho(X_{i_k}, X_{j_k} | \Theta(i_k, j_k)) - \rho)^2$$

$$\text{subject to : } E[\Delta_k | \delta_k, n_k] = \frac{(\gamma_k + 2)E[X|a_k, m_k, b_k] - 1}{\gamma_k}, k = 1, \ldots, n,$$

$$\delta_k = \frac{(n_k + 1)E[\Delta_k | \delta_k, n_k] - 1}{(n_k - 1)}, k = 1, \ldots, n,$$

$$V[X|\gamma_k, \delta_k, n_k] = V[X|a_k, m_k, b_k],\ k = 1, \ldots, n,$$

- $E[X|a_k, m_k, b_k]$ is the PERT Mean, $V[X|a_k, m_k, b_k]$ is the PERT variance.
- $\Theta(p, q) = (\gamma_p, \delta_p, n_p, \gamma_q, \delta_q, n_q), \mathcal{S} = (\gamma_k > \gamma_k^*, n_k > n_k^*, k = 1, \ldots n)$

Table 1. **Activity input parameters** (a_x, m_x, b_x) and **12 specified correlations** for the PERT network with resulting parameter solutions $(\gamma_x, n_x, \delta_x)$ that **match the PERT Mean and PERT Variance** and are solved for using **the constrained optimization problem above**.

#	X	a_x	m_x	b_x	Pairs	$\rho(X,Y)$	γ_x	n_x	δ_x	$x_{0.95}$
1	A	1	5	12	(A,B)	0.50	15.995	3.689	0.322	8.671
2	B	2	4	7	(B,F)	0.50	11.947	4.123	0.372	5.590
3	C	4	5	7	(C,E)	0.50	11.516	4.256	0.289	6.041
4	D	5	6	9	(D,G)	0.50	13.804	4.054	0.184	7.537
5	E	6	10	14	(D,H)	0.50	17.406	3.489	0.500	12.197
6	F	3	10	20	(E,J)	0.50	16.369	3.603	0.383	15.318
7	G	8	9	10	(F,J)	0.50	12.526	3.981	0.500	9.549
8	H	8	10	15	(G,K)	0.50	11.189	4.358	0.231	12.577
9	I	2	6	10	(H,I)	0.50	17.406	3.489	0.500	8.197
10	J	1	3	7	(H,K)	0.50	11.516	4.256	0.289	5.082
11	K	1	3	6	(I,J)	0.50	15.994	3.643	0.368	4.590
12	L	2	5	10	(J,L)	0.50	15.950	3.677	0.336	7.630

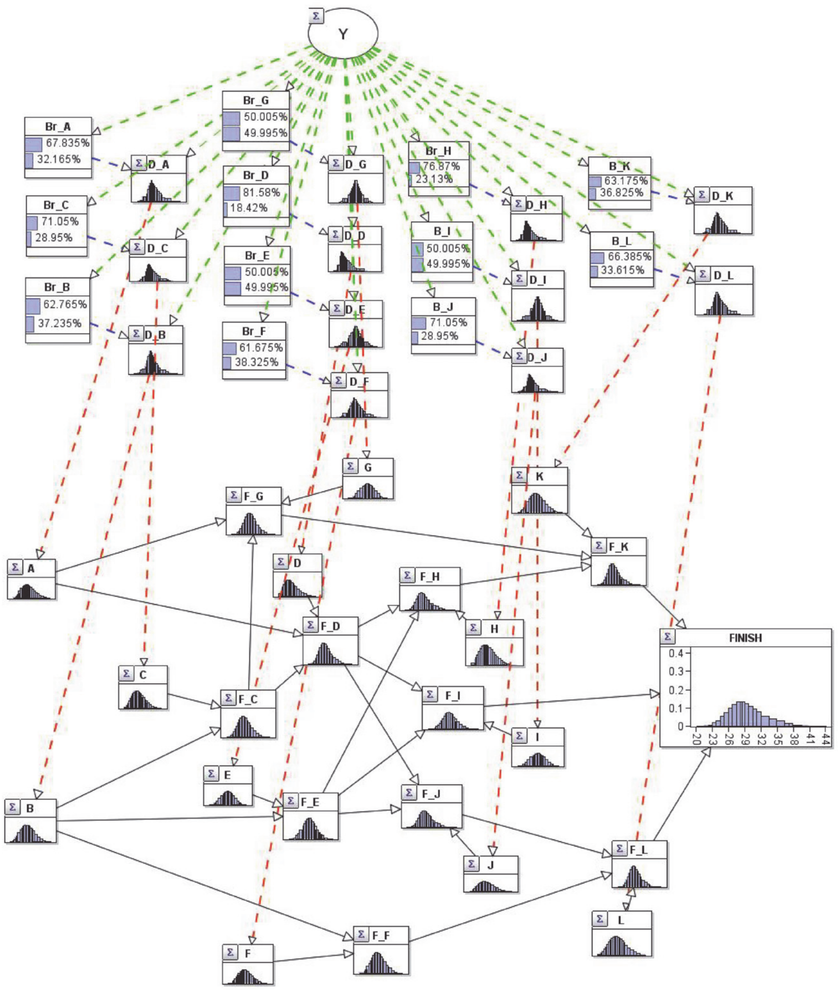
- When **using the 95-th percentiles $x_{0.95}$** for the activities, **the critical path has length of 36.34** and is indicated using the same **red arrows**.

Correlation Matrix for the 12 activities in 6-node project network

	A	B	C	D	E	F	G	H	I	J	K	L
A	**1.000**	0.500	0.502	0.562	0.573	0.571	0.501	0.505	0.573	0.502	0.568	0.570
B	0.500	**1.000**	0.438	0.489	0.504	0.500	0.440	0.441	0.504	0.438	0.497	0.499
C	0.502	0.438	**1.000**	0.494	0.500	0.500	0.437	0.445	0.500	0.441	0.497	0.500
D	0.562	0.489	0.494	**1.000**	0.554	0.558	0.500	0.500	0.554	0.494	0.555	0.559
E	0.573	0.504	0.500	0.554	**1.000**	0.576	0.514	0.500	0.588	0.500	0.572	0.572
F	0.571	0.500	0.500	0.558	0.576	**1.000**	0.504	0.502	0.576	0.500	0.568	0.569
G	0.501	0.440	0.437	0.500	0.514	0.504	**1.000**	0.437	0.514	0.437	0.500	0.500
H	0.505	0.441	0.445	0.500	0.500	0.502	0.437	**1.000**	0.500	0.445	0.500	0.503
I	0.573	0.504	0.500	0.554	0.588	0.576	0.514	0.500	**1.000**	0.500	0.572	0.572
J	0.502	0.438	0.441	0.494	0.500	0.500	0.437	0.445	0.500	**1.000**	0.497	0.500
K	0.568	0.497	0.497	0.555	0.572	0.568	0.500	0.500	0.572	0.497	**1.000**	0.566
L	0.570	0.499	0.500	0.559	0.572	0.569	0.500	0.503	0.572	0.500	0.566	**1.000**

- The correlations in **red** above the diagonal and in **green** below the diagonal equate to **the "project degree of dependence" $\rho = 0.5$**.
- **The Bayes Network dependence model** has **12 dependence parameters** γ_x, whereas the correlation matrix contains $\binom{12}{2} - 12 = 54$ **correlations.**
- Thus, **the Bayes Network dependence model** is **restricted in its flexibility** since, for example, the correlation $\rho(B, C) \approx 0.438$ **follows.**

Screenshot software **AgenaRisk®.**

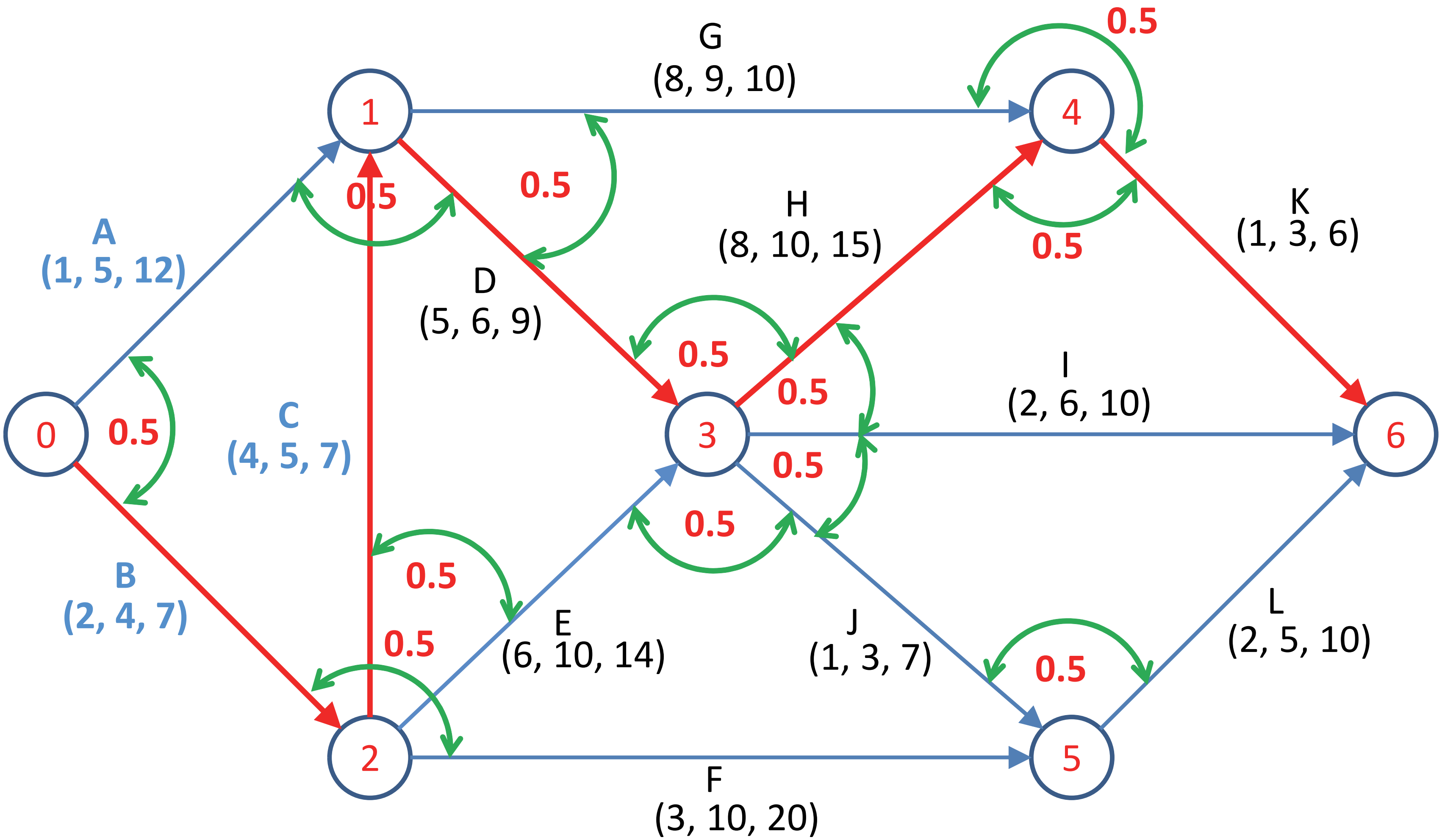


- **The green dashed arcs** lead from the common quantile node Y to the branch probability nodes and the mode relative distance nodes Δ_x.

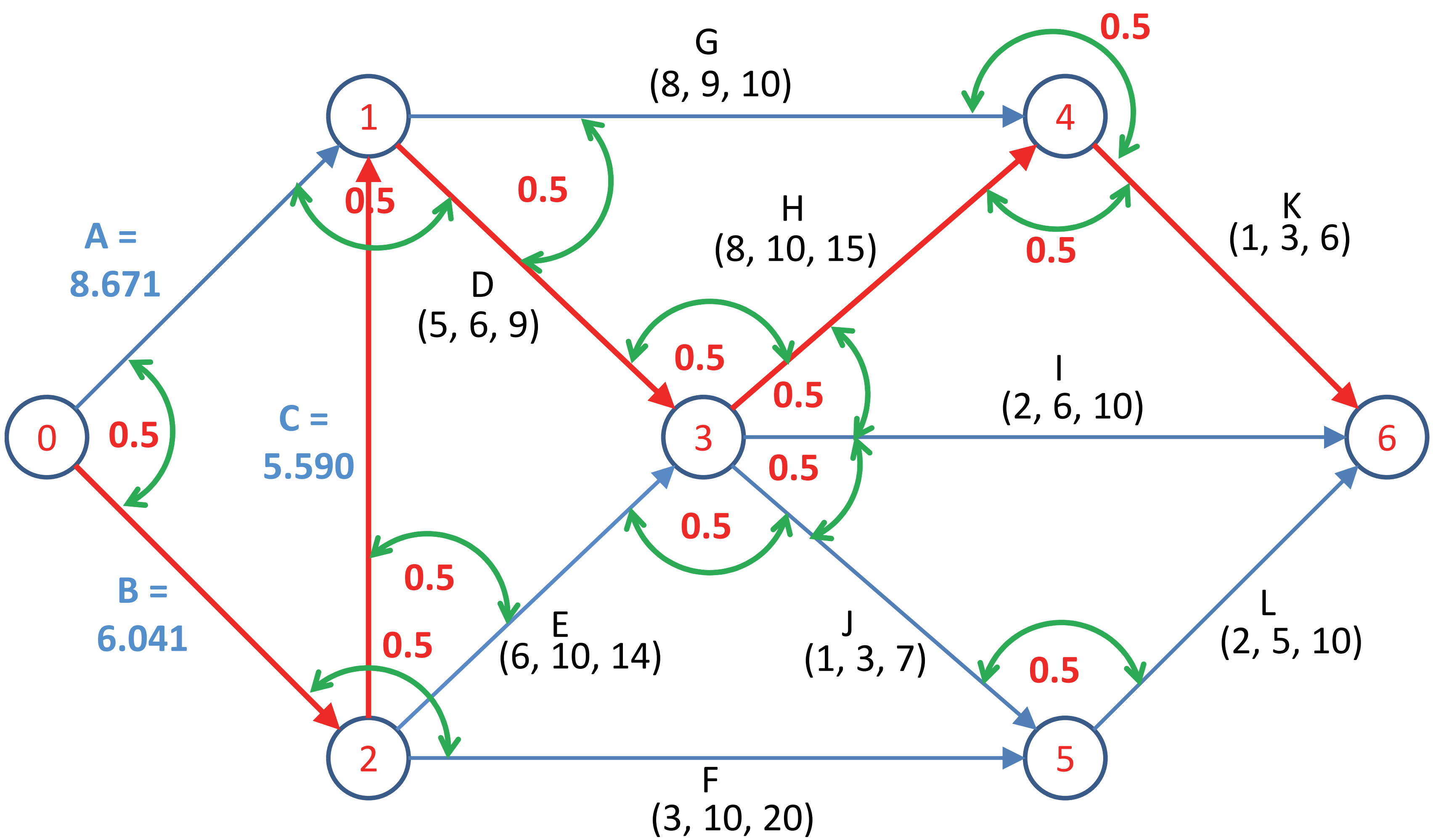
- **The red dashed arcs** lead from the mode relative distance nodes Δ_x to the activity duration nodes X.

- **The black solid arcs** in Figure 6 point to finishing time nodes "F_X" that contain the usual CPM expressions.

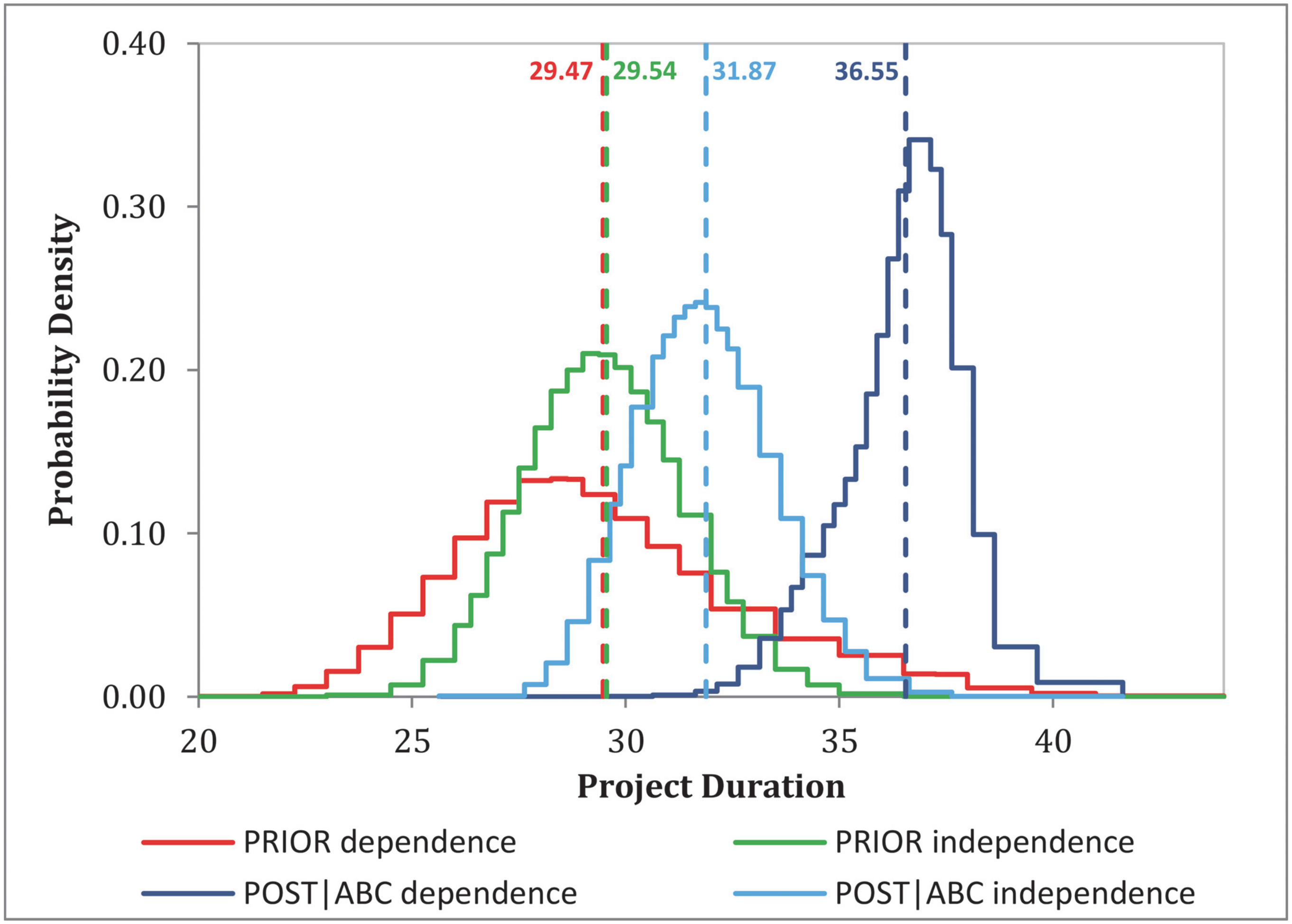
- Finally, **the "FINISH" node** in evaluates **the project completion time uncertainty.**



- Evaluate **prior project completion time distribution** under **dependence** and **independence.**



- Evaluate **posterior project completion time distr.** under **dependence** and **independence** given **completion A, B and C at their 95-th percentiles.**



Prior project completion time distribution results comparison for the project network and **posterior completion time distribution** comparison **given the completion of activities A, B and C at $a_{0.95}$, $b_{0.95}$ and $c_{0.95}$ quantile levels.**

- In comparing the prior project completion time distributions in red and green in, one observes **a larger uncertainty apriori in the statistical dependence case** (in red) which is **a known result** (see, e.g. Van Dorp, 2005).

- **The mean prior project completion** under both scenarios are evaluated at **≈ 29.47 and ≈ 29.54** for the dependence and independence scenario, respectively, which is slightly larger than **the critical path length 28 when using the most likely values m_x** for the activity durations.

- **Given the completion of the first three activities A, B and C in the project network in Figure 6 at their 95 percentile values $a_{0.95}$, $b_{0.95}$ and $c_{0.95}$**, the posterior mean project completion time **under statistical independence is evaluated by AgenaRisk® at 31.87** and **under statistical dependence evaluated at 36.55. The posterior mean of the common quantile node Y is evaluated by AgenaRisk® at ≈ 0.97.**

- It is worthwhile to note here that **the length of the critical path** using **all the 95 percentile values** was **evaluated at ≈ 36.34.**

- In other words, **the posterior mean project completion time forecast of 36.55**, given the completion times for activities A, B and C, is **more accurate under statistical dependence** than **the posterior mean project completion time forecast of 31.87 under statistical independence**.

- **Comparing the posterior project completion time distributions under statistical independence and under statistical dependence**, **a lesser uncertainty in the project completion time distribution under statistical dependence** whereas **the converse was true apriori.**

- While one also observes a reduction in uncertainty from the apriori statistical independence case to the aposteriori case, **the reduction in uncertainty is more pronounced in the statistical dependence case** going from the red distribution to the dark blue one.

OUTLINE

1. INTRODUCTION
2. ACTIVITY DURATION UNCERTAINTY MODEL
3. BAYES NETWORK STATISTICAL DEPENDENCE MODEL
4. BAYES NETWORK DEGREE OF DEPENDENCE ANALYSIS
5. DEPENDENT PERT EXAMPLE
6. **SUMMARY AND CONCLUSIONS**
7. SELECTED REFERENCES

6. SUMMARY AND CONCLUSIONS

- It would appear that **one forecasts more accurately in terms of the project completion time point estimate under statistical dependence**.

- **One also learns faster under statistical dependence** in terms of **a reduction of uncertainty in the project completion time distribution**.

- Heuristically and pragmatically, perhaps **a better updated project completion time point estimate can be obtained by averaging the quantile levels of the activities** that have completed and **utilizing that average quantile level for the remaining activities that have not completed in a single CPM evaluation step**, rather than **evaluating the remaining project completion time uncertainty and its point estimate under an assumption of statistical independence.**

- That being said, **to evaluate the remaining uncertainty in the project completion time under statistical dependence in a coherent manner following the completion of some activities, a Bayesian inference procedure is required**.

Gallagher, C. (1987). A note on PERT assumptions, *Management Science*, 33 (10), 1360.

Golenko-Ginzburg, D. (1988). On the distribution of activity time in PERT, *Journal of Operations Research,* 39, 767-771.

Hahn, E. D. (2008). Mixture densities for project management activity times: A robust approach to PERT. *European Journal of Operational Research*, 188(2), 450-459.

Herrerias-Pleguezuelo, R., Garcia-Perez, J. & Cruz-Rumbaud, S. (2003). A note on the reasonableness of PERT hypothesis, *Operations Research Letters*, 31 (1), 60-62

Herrerias-Velasco, J. M., Herrerias-Pleguezuelo, R., & Van Dorp, J. R. (2011). Revisiting the PERT mean and variance. *European Journal of Operational Research*, 210, 448-451.

Jenzarli, A. (1994). PERT belief networks, *Report 535*. The University of Tampa, FL.

Kamburowski, J. (1997). New validations of PERT times, *Omega,* 25 (3), 323-328.

Littlefield, T.J. & Randolph, P. (1987). An answer to Sasieni's question on PERT Times, *Management Science*, 33 (10), 1357-1359.

Malcolm, D., Roseboom, J., Clark, C., & Fazar, W. (1959). Application of a technique for research and development program application. *Operations Research*, 7 (5), 646-669.

Sasieni, M.W. (1986). A note on PERT Times, *Management Science*, 32 (12), 1652-1653.

Van Dorp, J. R., & Kotz, S. (2002). A Novel Extension of the Triangular Distribution and its Parameter Estimation. *The Statistician*, 51, 63-79.

Cho, S. (2009). A linear Bayesian stochastic approximation to update project duration estimates. *European Journal of Operations Research*, 196, 585-593.

Cho, S., & Covaliu, Z. (2003). Sequential estimation and crashing PERT networks with statistical dependence. *International Journal of Industrial Engineering*, 10, 391-399.

Covaliu, Z., & Soyer, R. (1997). Bayesian Learning in project management networks. *American Statistical Association Proceedings, Section on Bayesian Statistical Science*, (pp. 257-260).

Fang, C., & Marle, F. (2012). A simulation-based risk network model for decision support in project risk management. *Decision Support Systems*, 52, 635-644.

Nduka, I. & Van Dorp, J.R. (2015). Monitoring Uncertainty in Project Completion Times: A Bayesian Network Approach, *Unpublished Manuscript, https://www2.seas.gwu.edu/~dorpjr/Publications/JournalPapers/Nduka%20and%20Van%20Dorp%20-%20Unpublished%20Manuscript.pdf* , last accessed: 8/19/2018.

Neil, M., Tailor, M., & Marquez, D. (2007). Inference in Hybrid Bayesian Networks using Dynamic Discretization. *Statistics and Computing*, 17(3), 219-233.

Khodakarami, V., Fenton, N., & Neil, M. (2007). Project Scheduling: Improved Approach to Incorporate Uncertainty Using Bayesian Networks. *Project Management Journal*, 38, 39-49.

Van Dorp, J. R. (2005). Statistical Dependence through Common Risk Factors: With Applications in Uncertainty Analysis. *European Journal of Operational Research*, 161 (1), 240-255.

Van Dorp, J.R., & Duffey, M. (1999). Statistical dependence in risk analysis for project networks using Monte Carlo methods. *International Journal of Production Economics*, 58, 17-29.

Virto, M., Martin, J., & Insua, D. (2002). Approximate solutions of complex influence diagrams through MCMC methods. In S. Gamez (Ed.), *First European Workshop on Probabilistic Graphical Models*, (pp. 169-175).

7. REFERENCES... Software

Microsoft Project® (2018). Microsoft Project, https://en.wikipedia.org/wiki/ Microsoft_Project#Project_2016, last accessed: 8/19/2018.

Primavera® (2018). Primavera (Software), https://en.wikipedia.org/wiki/ Primavera_%28software%29, last accessed: 8/19/2018.

AgenaRisk® (2018). Agena, Bayesian Network and Simulation Software for Decision Support, http://www.agenarisk.com/, last accessed 8/19/2018.