Lad som sædvanlig $U(F_I) = p_d(a, \psi_{F_I}) - p_d(a, \psi_0)$ og $U_* = U(F_{\{1,\ldots,N\}})$. Som tidligere er faktorsvarene $\psi_{F_I}$ defineret således at det er det er de faktorsvar der giver den laveste $p_d(a, \psi)$ samtidig med at for $i \in I$ er $\psi_i = F_i$ - altså at vi har indsat de indtastede faktorsvar på alle index i $\psi$, $I$. Definer

$$S(F_I) = U(F_I) - \sum_{J \subseteq I, J \neq \emptyset, J \neq I} S(F_J)$$

Så gælder

$$U(F_I) = \sum_{J \subseteq I, J \neq \emptyset} S(F_i)$$

Dog er ikke alle $S(F_i)$'erne positive og kan derfor ikke bruges i dekompositionen. Dog påstår jeg

**Proposition 1.** *Definer*

$$V(F_i) = \sum_{J \subseteq \{1,\ldots,N\}, J \neq \emptyset, i \in J} \frac{S(F_J)}{|J|} \tag{1}$$

*for alle $i = 1, \ldots, N$. Da er*

$$V(F_i) \geq 0 \quad \forall i \in \{1, \ldots, N\} \tag{2}$$

$$\sum_{i=1}^{N} V(F_i) = U_* \tag{3}$$

*Bevis.* Beviset er ikke lavet, men det er understøttet af simulationer. □

Proposition 1 siger at $V(F_i)$'erne kan dekomponere $U_*$. Formen (1) kan fortolkes som at hver eneste faktor "arver" fra alle de interaktioner, hvor den indgår. Arven deles ligeligt mellem efterkommere.

**Lemma 1.** *Der gælder*

$$S(F_I) = \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|J|+|I|} U(F_J) \tag{4}$$

*Bevis.* Kan laves (og er lavet) med induktion. □

## 0.1 Multiplikativitet

I vores framework sker det ofte at der er multiplikativ effekt mellem riskratiotables.

$$p_d(a,\psi) = P_d(a) \cdot \frac{1}{\prod_{i=n}^{n} \text{mult}_i} \prod_{i=1}^{n} \text{RR}_i(\psi) \tag{5}$$

hvor $\text{RR}_i$ er en samlet riskratiotabel for den $i$'te riskfactor group. Definer nu

$$\mathcal{F}_i = \{\text{faktorer i den } i\text{'te riskfactor group}\} \tag{6}$$

$$U_i(F_I \cap \mathcal{F}_i) = \text{RR}_i(\psi_{F_I \cap \mathcal{F}_i}) \tag{7}$$

$$C = P_d(a) \cdot \frac{1}{\prod_{i=1}^{n} \text{mult}_i} \tag{8}$$

hvor $\psi_{F_I \cap \mathcal{F}_i}$ skal forstås som den $\psi$-vektor der giver den laveste værdi af $\text{RR}_i$-tabellen under betingelsen at $\psi$ har input-faktorsvarene for alle faktorer i $F_I$. (Som input i $\text{RR}_i$-tabellen er det lidt overflødigt at skrive $F_I \cap \mathcal{F}_i$ og ikke bare $F_I$, da $\text{RR}_i$ tabellen i forvejen ignorer alle faktorer der ikke har noget med den at gøre). Vi kan derfor skrive

$$p_d(a,\psi_{F_I}) = C \cdot \prod_{i=1}^{n} U_i(F_I \cap \mathcal{F}_i) \tag{9}$$

For at udnytte denne konstruktion bedst muligt ændrer vi også notation

$$\tilde{U}(F_I) := \frac{p_d(a,\psi_{F_I})}{C} = \prod_{i=1}^{n} U_i(F_I \cap \mathcal{F}_i)$$

$$\tilde{S}(F_I) = \tilde{U}(F_I) - \sum_{J \subseteq I, J \neq I} \tilde{S}(F_J)$$

Noter at

a. Nu eksisterer $\tilde{U}(F_\emptyset)$ og $\tilde{S}(F_\emptyset)$ og er trivielle

$$\tilde{U}(F_\emptyset) = \tilde{S}(F_\emptyset) = 1 \tag{10}$$

b. Forbindelsen mellem det gamle $U$ og det nye $\tilde{U}$ er

$$U(F_I) = C \cdot \left[\tilde{U}(F_I) - \tilde{U}(F_\emptyset)\right] \tag{11}$$

c. Sammenhængen mellem $\tilde{S}$ og $S$ er

$$S(F_I) = C \cdot \tilde{S}(F_I) \quad I \neq \emptyset \tag{12}$$

Sammenhængen (12) retfærdiggør også definitionen

$$\tilde{V}(F_i) = \sum_{J \subseteq \{1,\dots,N\}, J \neq \emptyset, i \in J} \frac{\tilde{S}(F_J)}{|J|}, \tag{13}$$

fordi med dette $\tilde{V}(F_i)$ kan vi nemt finde $V(F_i) = C \cdot \tilde{V}(F_i)$. Vi kan også lave en tilde-version af Lemma 1

**Korollar 1.** *Der gælder*

$$\tilde{S}(F_I) = \sum_{J \subseteq I} (-1)^{|J|+|I|} \tilde{U}(F_J) \tag{14}$$

*Bevis.* Det kan udregnes direkte

$$\begin{aligned}
\tilde{S}(F_I) &= \frac{1}{C} \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|I|+|J|} U(F_J) \\
&= \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|I|+|J|} \left[ \tilde{U}(F_J) - \tilde{U}(F_\emptyset) \right] \\
&= \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|I|+|J|} \tilde{U}(F_J) - (-1)^{|I|} U(F_\emptyset) \sum_{J \subseteq I, I \neq \emptyset} (-1)^{|J|} \\
&= \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|I|+|J|} \tilde{U}(F_J) - (-1)^{|I|} U(F_\emptyset) \sum_{s=1}^{|I|} (-1)^{|J|} \binom{|I|}{s} \\
&= \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|I|+|J|} \tilde{U}(F_J) \\
&\quad - (-1)^{|I|} U(F_\emptyset) \Bigg( \Bigg[ \sum_{s=0}^{|I|} (-1)^{|J|} \binom{|I|}{s} 1^{|I|-|J|} \Bigg] - 1 \Bigg) \\
&= \sum_{J \subseteq I, J \neq \emptyset} (-1)^{|I|+|J|} \tilde{U}(F_J) - (-1)^{|I|} U(F_\emptyset) \big( -1 \big) \\
&= \sum_{J \subseteq I} (-1)^{|I|+|J|} \tilde{U}(F_J)
\end{aligned}$$

□

**Lemma 2.** *Definer rekursivt*

$$S_i(F_I \cap \mathcal{F}_i) = \tilde{U}_i(F_I \cap \mathcal{F}_i) - \sum_{J \subseteq \cap \mathcal{F}_i, J \neq I} \tilde{S}_i(F_J) \tag{15}$$

*Så gælder*

$$S_i(F_I) = \sum_{J \subseteq I} (-1)^{|J|+|I|} \tilde{U}_i(F_J), \quad I \subseteq \mathcal{F}_i \tag{16}$$

*og*

$$S(F_I) = \prod_{i=1}^{n} S(F_{I \cap \mathcal{F}_i}) \tag{17}$$

*Bevis.* Formel (16) kan udledes på samme måde som et specialtilfælde af Korollar 1. Vi viser (17) for $n = 2$ og deraf følger resten nemt. Lad $I \subseteq \{1, \ldots, N\}$. Skriv $L = I \cap \mathcal{F}_1$, $K = I \cap \mathcal{F}_2$. Så er

$$\begin{aligned}
S(F_K)S(F_L) &= \sum_{J \subseteq K} (-1)^{|J|+|K|} \tilde{U}_1(F_J) \sum_{M \subseteq L} (-1)^{|M|+|L|} \tilde{U}_2(F_M) \\
&= \sum_{J \subseteq K} \sum_{M \subseteq L} (-1)^{|J|+|K|+|L|+|M|} \tilde{U}(F_{J \cup M}) \\
&= \sum_{J \subseteq K} \sum_{M \subseteq L} (-1)^{|J|+|K|+|L|+|M|} \tilde{U}(F_{J \cup M}) \\
&= \sum_{J \subseteq K \cup L} (-1)^{|J|+|L|+|K|} \tilde{U}(F_J) \\
&= \sum_{J \subseteq I} (-1)^{|J|+|I|} \tilde{U}(F_J) \\
&= S(F_I)
\end{aligned}$$

□

Med dette i hånden kan vi udregne $\tilde{V}(F_i)$ hurtigere.

**Sætning 1.** *Lad $i \in \{1, \ldots, N\}$ og lad $j_i$ være det indeks som opfylder $i \in \mathcal{F}_{j_i}$. Da er*

$$\tilde{V}(F_i) = \sum_{s=1}^{N} \frac{1}{s} \sum_{t=1}^{N-s} \alpha_{t,j_i,i} \beta_{t-s,\{1,\ldots,N\} \setminus \{j_i\}} \tag{18}$$

*hvor*

$$\alpha_{t,j,i} = \sum_{\substack{J \subseteq \mathcal{F}_j, i \in J \\ |J|=t}} S_j(F_J) \tag{19}$$

$$\beta_{t,j} = \sum_{\substack{J \subseteq \mathcal{F}_j \\ |J|=t}} S_j(F_J) \tag{20}$$

$$\beta_{t,I} = \sum_{\substack{t_i \in \{0,\ldots,N\} \\ ,i \in I, \sum t_i = t}} \left( \prod_{i \in I} \beta_{t_i,i} \right) \tag{21}$$

*Bevis.* Formlen følger ved at anvende Lemma 2 på $\tilde{V}(F_i)$. Antag for simplicitet at $j_i = 1$.

$$
\begin{aligned}
\tilde{V}(F_i) &= \sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset, i\in J}} \frac{\tilde{S}(F_J)}{|J|} \\
&= \sum_{s=1}^{N} \left( \sum_{\substack{s_1,\dots,s_n\in\{0,\dots,N\},\\ s_1+\cdots+s_n=s}} \left( \sum_{\substack{J_1\subseteq\mathcal{F}_1,\dots,J_n\subseteq\mathcal{F}_n\\ i\in J_{j_i}, |J_j|=s_j}} \frac{\tilde{\prod}_{j=1}^{n} S_j(F_{J_j})}{|s|} \right)\right) \\
&= \sum_{s=1}^{N} \frac{1}{s} \left( \sum_{\substack{s_1,\dots,s_n\in\{0,\dots,N\},\\ s_1+\cdots+s_n=s}} \left( \sum_{\substack{J_1\subseteq\mathcal{F}_1,\\ |J_1|=s_1, i\in J_1}} S_1(F_{J_1}) \right.\right. \qquad (22) \\
&\qquad \left.\left. \cdot \prod_{j=2}^{n} \sum_{J_j\subseteq\mathcal{F}_j, |J_j|=s_j} S_j(F_{J_j}) \right)\right) \\
&= \sum_{s=1}^{N} \frac{1}{s} \left( \sum_{\substack{s_1,\dots,s_n\in\{0,\dots,N\},\\ s_1+\cdots+s_n=s}} \left( \alpha_{s_1,1,i} \prod_{j=2}^{n} \beta_{s_j,j} \right)\right)
\end{aligned}
$$

Det ses at $\alpha_{0,i,j} = 0$ ud fra definitionen hvilket gør at vi kan nøjes med at summe over $s_1 \geq 0$. Vi kan også splitte summen over $s_1, \dots, s_n$ op i to; en over $s_1$ og en over de resterende variable. Det giver

$$
\begin{aligned}
V(F_i) &= \sum_{s=1}^{N} \frac{1}{s} \sum_{s_1=1}^{n} \alpha_{s_1,1,i} \left( \sum_{\substack{s_2,\dots,s_n\in\{0,\dots,s-s_1\},\\ s_2+\cdots+s_n=s-s_1}} \left( \prod_{j=2}^{n} \beta_{s_j,j} \right)\right) \\
&= \sum_{s=1}^{N} \frac{1}{s} \sum_{s_1=1}^{n} \alpha_{s_1,1,i} \beta_{s-s_1,\{1,\dots,N\}\setminus\{1\}}
\end{aligned}
$$

Dette viser sætningen for $j_1 = 1$. For et generelt $j_1 \in \{1, \dots, n\}$ er det bare summen $\sum S_{j_i}(F_{J_{j_i}})$, som skal trækkes uden for parentes i (22) og så følger den generelle sætning på lignende vis. □

## 0.2 Vægtning af faktorer

Der er stor forskel på hvor stor forskel der er på hvor nemt det er at ændre en risikofaktor. Ens biologiske køn kan man for eksempel ikke ændre, men man kan godt variere hvor meget man ryger. Vi forventer derfor at brugere er mindre interesserede i interaktionen mellem køn og rygning og mere interesseret i hvad deres forskellige rygevaner gør *givet deres køn*. Derfor forestiller vi os at alle faktorer får en diskret vægtning afhængig af hvor nemme de er at ændre. Lad $F_{Q_1}$ være de faktorer, der er vægtet lavest. Da dekomponerer vi først

$$\tilde{U}(F_{Q_1}) \propto p_d(a, \psi_{F_{Q_1}}) \tag{23}$$

med $U(F_\emptyset) = p_d(a, \psi_0)$. Hvis der er faktorer med en anden rangering samlet i $F_{Q_2}$, komponerer vi dernæst

$$\tilde{U}(F_{Q_2} \cup F_{Q_1}) \propto p_d(a, \psi_{F_{Q_2 \cup Q_1}}) \tag{24}$$

med $U(F_\emptyset) \propto p_d(a, \psi_{F_{Q_1}})$. Det vil sige at vi lader som om $\psi_0$ har fået indsat de indtastede $F_{Q_1}$-værdier. Som konsekvens bliver alle eventuelle forstærkende interaktioner mellem $Q_1$-variablene og $Q_2$-variablene lavet om til rene $Q_2$-bidrag. Omvendt bliver eventuelle inhiberende interaktioner lavet om til rene $Q_1$-bidrag. Denne virkemåde er lige, hvad vi vil have! Hvis der er flere distinkte niveauer af rangeringer, sker udregningen på samme måde som i (24).

Ligesom før kan vi bare ignorere normaliseringskonstanten $C$ og først gange den på til sidst, da det er lige meget om vi dekomponerer $p_d(a, \psi_{F_Q})$ eller $p_d(a, \psi_{F_Q})/C$. Vi skal dog, selvfølgelig, huske at udregne det fulde $\tilde{U}$ baseret på produktet af alle $U_i$'erne.

$$\tilde{U}(F_Q) = \prod_{i=1}^{n} U_i(F_Q \cap \mathcal{F}_i) \tag{25}$$

## 0.3 Algoritme

Baseret på Sætning 1 og ovenstående sektion, foreslås Algorithm 1.

## 0.4 Interaktioner

I stedet for at reducere al skyld til kun en enkelt faktor kan vi også formidle mere af den information, der ligger i $S(F_J)$'erne. Det store problem er at nogle af $S(F_J)$'erne kan være negative og så kan de ikke visualiseres på en nem måde. Jeg vil her foreslå et kompromis mellem at formidle $S(F_J)$'erne i deres helhed og på en forståelig måde.

### 0.4.1 Containment

I dette scheme går vi efter at afgrænse de negative interaktioner og tage dem som et specialtilfælde. Vi betrager kun en enkelt riskfactorgruppe, $\mathcal{F} =$

**Algorithm 1** Udregning af $V(F_i)$'er

**procedure** GETINNERCAUSES(f)<br>
    $F_{Q_1}, \dots, F_{Q_r} \leftarrow \text{divideFactorsBasedOnWeight}(F_{\{1,\dots,N\}})$<br>
    $\mathit{InnerCauses} \leftarrow \text{list}()$<br>
    **for** $F_Q \in \{F_{Q_1}, \dots, F_{Q_r}\}$ **do**<br>
        $\beta_{s,j} \leftarrow 0$ for all $s, j$<br>
        $\alpha_{s,j,i} \leftarrow 0$ for all $s, j, i$<br>
        **for** $j = 1, \dots, n$ **do**<br>
            **for** $J \subseteq F_Q \cap \mathcal{F}_j$ **do**<br>
                compute $S_j(F_J)$<br>
                $s \leftarrow |J|$<br>
                $\beta_{s,j} \leftarrow \beta_{s,j} + S_j(F_J)$<br>
                **for** $i \in F_Q \cap \mathcal{F}_j$ **do**<br>
                    $\alpha_{s,j,i} \leftarrow \alpha_{s,j,i} + S_j(F_J)$<br>
        $(\beta_{s,\{1,\dots,n\}\setminus\{j\}})_{s\geq 0, j=1,\dots,n} \leftarrow \text{computeBetaSums}(\{\beta_{s,j}\})$<br>
        $V(F_i) \leftarrow 0 \;\text{ for all }\; i \in F_Q$<br>
        **for** $i \in F_R$ **do**<br>
            **for** $k = 1, \dots, |\mathcal{F}_{j_i}|$ **do**<br>
                **for** $l = 0, \dots, |F_Q| - |\mathcal{F}_{j_i}|$ **do**<br>
                    $V(F_i) \leftarrow \alpha_{k,j_i,i}\beta_{l,\{1,\dots,n\}\setminus\{j_i\}}/(k+l)$<br>
        $\mathit{InnerCauses}.\text{add}\big((V(F_i))_{i\in F_Q}\big)$<br>
    **return** $\mathit{InnerCauses}$

$\{1, \dots, M\}$. Lad

$$\begin{aligned}\mathcal{H}_i = \big\{\, k \in \mathcal{F} \mid \exists m : J_1, \dots, J_m \subseteq \mathcal{F}, i \in J_1, k \in J_m, J_j \neq \emptyset, \\ S(F_{J_1 \cup J_2}) < 0, \dots, S(F_{J_{m-1} \cup J_m}) < 0 \big\}\end{aligned} \tag{26}$$

Med andre ord er $\mathcal{H}_i$ alle de faktorer, som den $i$'te faktor har en negativ interaktion med - eller alle de faktorer som er forbundne til den $i$'te faktor via negative interaktioner. Mængderne $\mathcal{H}_1, \dots, \mathcal{H}_M$ udgør en inddeling af $F$, hvis vi ser bort fra duplikaterne. For alle (unikke) $\mathcal{H} \in \{\mathcal{H}_1, \dots, \mathcal{H}_M | H_j \neq \emptyset\}$ gør vi det følgende:

1. Udregn
$$W(F_{\mathcal{H}}) = U(F_{\mathcal{H}}) - \sum_{i \in \mathcal{H}} S(F_i) \tag{27}$$
Man skal fortolke $W(F_{\mathcal{H}})$ som forskellen mellem den samlede sandsynlighed i $F_{\mathcal{H}}$ og den der kan forklares udelukkende fra de marginale. Hvis interaktionen mellem faktorerne i $\mathcal{H}$ er additiv ville $W(F_{\mathcal{H}}) = 0$. Da vi ved der er negative interaktioner i $\mathcal{H}$ vil vi dog ofte forvente at $W(F_{\mathcal{H}}) < 0$, men det kan også ske at $W(F_{\mathcal{H}}) \geq 0$.

2. Hvis $W(F_{\mathcal{H}}) \geq 0$ laver vi dekomponeringen
$$F_i \mapsto S(F_i) \quad \forall i \in \mathcal{H} \tag{28}$$
$$F_{\mathcal{H}} \mapsto W(F_{\mathcal{H}}) \tag{29}$$
og alle andre faktormængder $I \subseteq H$ får værdien 0. Da vil
   - Summen af alle led i dekomponeringen give $U(F_{\mathcal{H}})$ som ønsket
   - Alle led vil være positive
   - Fortolkningen af leddet hørende til $F_{\mathcal{H}}$ vil være lidt generisk som "en uspecificeret multifaktoriel interaktion mellem $F_{\mathcal{H}}$".

3. Hvis $W(F_{\mathcal{H}}) < 0$ vil vi transformere noget af eller hele $S(F_i)$-summen om til ELLER-led. Definer hertil $S^{(i)}$ som er det $i$'te mindste tal i mængden $\{S(F_i) \mid i \in \mathcal{H}\}$
   (i) Hvis $-W(F_{\mathcal{H}}) \leq S^{(1)} \frac{|\mathcal{H}|}{2}$ kan vi lave dekomponeringen
$$F_i \mapsto S(F_i) + W(F_{\mathcal{H}}) \frac{2}{|\mathcal{H}|} \quad \forall i \in \mathcal{H} \tag{30}$$
$$F_{\mathcal{H}} \mapsto -W(F_{\mathcal{H}}) \tag{31}$$

Vi vil igen have at all leddene er positive og vil summe til $U(F_{\mathcal{H}})$. Fortolkningen af $F_{\mathcal{H}}$-leddet vil være

$$F_{j_1} \text{ ELLER } F_{j_2} \text{ ELLER } \cdots \text{ ELLER } F_{j_{|\mathcal{H}|}} \tag{32}$$

for $\mathcal{H} = \{j_1, \ldots, j_{|\mathcal{H}|}\}$. Dette ELLER-led skal forstås som at risikoen er uniformt fordelt henover dets medlemmer.

(ii) Hvis $-W(F_{\mathcal{H}}) > S^{(1)}\frac{|\mathcal{H}|}{2}$ trækker vi først $S^{(|\mathcal{H}|)}\frac{2}{|\mathcal{H}|}$ fra alle de marginale $S(F_i)$'er og fordeler derefter resten af det negative $W(F_{\mathcal{H}})$ blandt de resterende positive $S(F_i)$'er. Det fortsættes med indtil al $W(F_{\mathcal{H}})$-"gælden" er fordelt. Hvis ikke vi når at uddele al gælden før vi løber tør for $S(F_i)$-masse, reskalerer vi til sidst. Hvis vi definerer dekomponeringsleddet for $F_I$ som $D(F_I)$, kan det beskrives algoritmisk ved Algorithm 2

**Algorithm 2** Udregning af $D(F_I)$'er

**procedure** $\textsc{DistributeDebt}(\{S(F_i)\}_{i\in\mathcal{H}}, W(F_{\mathcal{H}}))$
  $D(F_i) \leftarrow S(F_i)$ for all $i \in \mathcal{H}$.
  $D(F_I) \leftarrow 0$ for $I \subseteq \mathcal{H}$.
  debt $\leftarrow -W(F_{\mathcal{H}})$.
  $R_1, \ldots, R_{|\mathcal{H}|} \leftarrow \text{ranks}(\{S(F_i)\}_{i\in\mathcal{H}}, \text{ties} = \text{“assign at random”})$
  $S^{(i)} \leftarrow S(F_{R_i})$ for all $i \in \mathcal{H}$
  $j \leftarrow 1$
  **while** debt $> 0$ **and** $j \leq |\mathcal{H}|$ **do**
    debt_reduction $\leftarrow \min(\text{debt}, (S^{(j)} - S^{(j-1)}) \cdot (|H| - j + 1)/2)$
    **for** $k = j, \ldots, |\mathcal{H}|$ **do**
      $D(F_{R_k}) \leftarrow D(F_{R_k}) - \text{debt\_reduction} \cdot 2/(|H| - j + 1)$
    $D(F_{\mathcal{H}\setminus\{R_1,\ldots,R_{j-1}\}}) \leftarrow$ debt_reduction
  **if** debt $> 0$ **then**
    rescale $\leftarrow (W(F_{\mathcal{H}}) - \sum_{i\in\mathcal{H}} S(F_i)) / \sum_{I\subseteq\mathcal{H}} D(F_I)$
    $D(F_I) \leftarrow D(F_I) \cdot$ rescale for alle $I \subseteq \mathcal{H}$.
  **return** $\{D(F_I)\}_{I\subseteq\mathcal{H}}$

Efter man har kørt ovenstående er alle de resterende interaktioner inden for riskfactorgruppen positive og vi kan altså bruge dekomponeringen

$$U(F_{\mathcal{F}}) = \sum_{I\subseteq\mathcal{F}} E(F_I) \tag{33}$$

$$E(F_I) = \begin{cases} D(F_I) & \text{hvis der er et } i \text{ så } I \subseteq \mathcal{H}_i \\ S(F_I) & \text{ellers.} \end{cases} \tag{34}$$

Når der er $n$ riskfactorgrupper ganger vi de enkelte led sammen

$$E(F_J) = \prod_{i=1}^{n} E_i(F_J \cap \mathcal{F}_i) \tag{35}$$

Hvis $E_i(F_J \cap \mathcal{F}_i) = S_i(F_J \cap \mathcal{F}_i)$ for alle $i$ vil $E(F_J) = S(F_J)$ på grund af Lemma 2. Vi kan da fortolke $S(F_J)$ hvor $J = \{j_1, \dots, j_{|J|}\}$ som

$$\text{at dø af dødsårsagen på grund af } F_{j_1} \textbf{ og } F_{j_2} \textbf{ og } \cdots \textbf{ og } F_{j_{|J|}} \tag{36}$$

Hvis der i (35) findes led hvor $E_i(F_J \cap \mathcal{F}_i) = D_i(F_J \cap \mathcal{F}_i)$ er der ikke umiddelbart en mening med produktet. Jeg foreslår dog denne fortolkning

$$\text{Fortolkning}(E_1(J \cap \mathcal{F}_1)) \textbf{ og } \cdots \textbf{ og } \text{Fortolkning}(E_n(J \cap \mathcal{F}_n)), \tag{37}$$

$$\text{Fortolkning}(E_i(K) = \begin{cases} F_{k_1} \textbf{ eller} \cdots \textbf{eller } F_{k_{|K|}} & \text{hvis } D_i(F_K) = E_i(F_K) \\ F_{k_1} \textbf{ og} \cdots \textbf{og } F_{k_{|K|}} & \text{hvis } S_i(F_K) = E_i(F_K) \end{cases} \tag{38}$$

Det bibeholder og-fortolkningen mellem led fra forskellige risk factorgrupper.

# A Tør ikke slette endnu

**Lemma 3.** *Vi kan skrive*

$$V(F_i) = \sum_{K \subseteq \{1,\dots,N\}, K \neq \emptyset} U(F_K) a_K \tag{39}$$

*hvor*

$$a_K = \begin{cases} \sum_{s=0}^{N-|K|} (-1)^s \frac{1}{|K|+s} \binom{N-|K|}{s} & \text{hvis } i \in K \\ \sum_{s=1}^{N-|K|} \frac{1}{|K|+s} \binom{N-|K|-1}{s-1} (-1)^s & \text{hvis } i \notin K \end{cases} \tag{40}$$

*Bevis.* Formlen kan ses ved at skrive

$$\begin{aligned} V(F_i) &= \sum_{J \subseteq \{1,\dots,N\}, J \neq \emptyset, i \in J} \frac{S(F_J)}{|J|} \\ &= \sum_{J \subseteq \{1,\dots,N\}, J \neq \emptyset, i \in J} \frac{1}{|J|} \sum_{K \subseteq J, K \neq \emptyset} (-1)^{|J|+|K|} U(F_K) \\ &= \sum_{K \subseteq \{1,\dots,N\}, K \neq \emptyset} \left( U(F_K) \sum_{J, K \subseteq J, i \in J} (-1)^{|J|+|K|} \frac{1}{|J|} \right) \end{aligned}$$

Vi kan derfor konkludere at $a_K$ findes og er lig med

$$a_K = \sum_{J, K \subseteq J, i \in J} (-1)^{|J|+|K|} \frac{1}{|J|} \tag{41}$$

Det kan vi skrive om ved at stratificere efter $|J|$.

$$\begin{aligned} a_K &= \sum_{s=0}^{N} \sum_{J, K \subseteq J, i \in J, |J|=s} (-1)^{s+|K|} \frac{1}{s} \\ a_K &= \sum_{s=0}^{N} (-1)^{s+|K|} \frac{1}{s} \# \big\{ J, K \subseteq J, i \in J, |J| = s \big\} \end{aligned}$$

Mængden $\big\{J, K \subseteq J, i \in J, |J| = s\big\}$ er tom hvis $s < |K|$. Den er også tom hvis $i \notin K$ og $|s| = K$. Hvis $i \in K$ og $s \geq |K|$ kan vi vælge elementerne

$J \setminus K$ uniformt fra $\{1, \ldots, N\} \setminus K$. Det giver $\binom{N-|K|}{s-|K|}$ muligheder. Dvs.

$$\begin{aligned}
a_K &= \sum_{s=|K|}^{N} (-1)^{s+|K|} \frac{1}{s} \binom{N-|K|}{s-|K|} \\
&= \sum_{s=|K|}^{N} (-1)^{s-|K|} \frac{1}{s} \binom{N-|K|}{s-|K|} \\
&= \sum_{s=0}^{N-|K|} (-1)^{s} \frac{1}{s+|K|} \binom{N-|K|}{s}
\end{aligned}$$

Hvis $i \notin K$, består alle elmenter i $\big\{J, K \subseteq J, i \in J, |J| = s\big\}$ af foreningen af $K$, $\{i\}$ og et antal indices fra $\{1, \ldots, N\} \setminus (K \cup \{i\})$. Disse indices skal igen vælges uniformt, og hvis der skal være $s$ elementer i $J$, kan det gøres på $\binom{N-|K|-1}{s-|K|-1}$ måder. Alt i alt får vi

$$\begin{aligned}
a_K &= \sum_{s=|K|+1}^{N} (-1)^{s+|K|} \frac{1}{s} \binom{N-|K|-1}{s-|K|-1} \\
&= \sum_{s=1}^{N-|K|} (-1)^{s} \frac{1}{s+|K|} \binom{N-|K|-1}{s-1}
\end{aligned}$$

$\square$

Vi kan også udlede en tilde-version af Lemma 3. Hvis vi følger udregningen

$$\begin{aligned}
\tilde{V}(F_i) &= \sum_{J \subseteq \{1,\ldots,N\}, J \neq \emptyset, i \in J} \frac{\tilde{S}(F_J)}{|J|} \\
&= \sum_{J \subseteq \{1,\ldots,N\}, J \neq \emptyset, i \in J} \frac{1}{|J|} \sum_{K \subseteq J} (-1)^{|J|+|K|} \tilde{U}(F_K) \\
&= \sum_{K \subseteq \{1,\ldots,N\}} \left( \tilde{U}(F_K) \sum_{J, K \subseteq J, i \in J} (-1)^{|J|+|K|} \frac{1}{|J|} \right)
\end{aligned} \tag{42}$$

Vi ser straks at koefficienterne for $\tilde{U}(F_K), K \neq \emptyset$ er de samme som i Lemma 3. Lad os nu betrage $a_\emptyset$. Som i beviset for Lemma 3 kan vi se at koefficienten for $a_\emptyset$ er

$$
\begin{aligned}
a_\emptyset &= \sum_{s=1}^{|N|} \frac{1}{s}(-1)^s \#\{J, J \subseteq \emptyset, i \in J, |J| = s\} \\
&= \sum_{s=1}^{|N|} \frac{1}{s}(-1)^s \#\{J, i \in J, |J| = s\}
\end{aligned}
$$

Et $J$ i mængden $\{J, i \in J, |J| = s\}$ kan vælges på $\binom{N-1}{s-1}$ måder, da der er så mange forskellige valg af elementer som ikke er $i$. Det giver

$$
a_\emptyset = \sum_{s=1}^{|N|} \frac{1}{s}(-1)^s \binom{N-1}{s-1} \tag{43}
$$

Alt i alt kan vi konkludere følgende korollar

**Korollar 2.** *Vi kan skrive*

$$
\tilde{V}(F_i) = \sum_{K \subseteq \{1,\dots,N\}} \tilde{U}(F_K) a_K \tag{44}
$$

*hvor*

$$
a_K = \begin{cases} \sum_{s=0}^{N-|K|} (-1)^s \frac{1}{|K|+s} \binom{N-|K|}{s} & \text{hvis } i \in K \\ \sum_{s=1}^{N-|K|} \frac{1}{|K|+s} \binom{N-|K|-1}{s-1} (-1)^s & \text{hvis } i \notin K \end{cases} \tag{45}
$$

Udregningen af (44) er umiddelbart af kompleksitet $O(2^N)$ hvor $N$ er antallet af faktorer. Det kan godt være lang tid hvis der er mange riskfactor grupper, så vi ønsker at finde en hurtigere måde at udregne det på. Antag som før at der er $n$ grupper. Så kan vi skrive

$$
\tilde{V}(F_i) = \sum_{K \subseteq \{1,\dots,N\}} a_K \prod_{j=1}^{n} U_j(F_K \cap \mathcal{F}_j)
$$

Vi kan nu bruge at $a$ kun afhænger af $K$ gennem $|K|$ samt sandhedsværdien af $i \in K$. Notationsmæssigt indfører vi derfor $a_s(i \in K)$ og $a_s(i \notin K)$, som står for $a_K$-koefficienten når $|K| = s$ og $i$ hhv. ligger og ikke ligger i $K$.

$$
\begin{aligned}
\tilde{V}(F_i) = \sum_{s=0}^{N} \Bigg( & a_s(i \in K) \sum_{\substack{K \subseteq \{1,\dots,N\}, \\ |K|=s, i \in K}} \prod_{j=1}^{n} \tilde{U}_j(F_K \cap \mathcal{F}_j) \\
& + a_s(i \notin K) \sum_{\substack{K \subseteq \{1,\dots,N\}, \\ |K|=s, i \notin K}} \prod_{j=1}^{n} \tilde{U}_j(F_K \cap \mathcal{F}_j) \Bigg)
\end{aligned} \tag{46}
$$

Lad nu $j(i)$ være den af de $n$ riskfactor groups hvor $i$ ligger. Antag uden tab af generalitet at $j(i) = 1$. På grund af vores riskfactor struktur har vi sammenhængen $\{1, \ldots, N\} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n$, hvilket vi kan bruge til at skrive

$$\sum_{\substack{K \subseteq \{1,\ldots,N\},\\ |K|=s, i \in K}} \prod_{j=1}^{n} \tilde{U}_j(F_K \cap \mathcal{F}_j) \tag{47}$$

$$= \sum_{\substack{s_1,\ldots,s_n \in \{0,\ldots,N\},\\ s_1+\cdots+s_n=s}} \sum_{\substack{K_1 \subseteq \mathcal{F}_1,\\ |K_1|=s_1, i \in K_1}} \sum_{\substack{K_2 \subseteq \mathcal{F}_2,\\ |K_2|=s_2}} \cdots \sum_{\substack{K_n \subseteq \mathcal{F}_n,\\ |K_n|=s_n}} \prod_{j=1}^{n} \tilde{U}_j(F_{K_j} \cap \mathcal{F}_j) \tag{48}$$

$$= \sum_{s_1 \in \{0,\ldots,N\}} \sum_{\substack{K_1 \subseteq \mathcal{F}_1,\\ |K_1|=s_1, i \in K_1}} \tilde{U}_1(F_{K_1}) \sum_{s_2 \in \{0,\ldots,N\}} \sum_{\substack{K_2 \subseteq \mathcal{F}_2,\\ |K_2|=s_2}} \tilde{U}_2(F_{K_2}) \cdots \tag{49}$$

$$\cdots \sum_{\substack{K_n \subseteq \mathcal{F}_2,\\ |K_n|=s-s_1-\cdots-s_{n-1}}} \tilde{U}_n(F_{K_n})$$

Dette udtryk er meget nyttigt da det i høj grad kan splittes op i produkter af summer. Lad for alle $s = 0, \ldots, N$

$$\beta_{s,j} = \sum_{\substack{K \subseteq \mathcal{F}_j,\\ |K|=s}} \tilde{U}_j(F_K), \quad j = 1, \ldots, n \tag{50}$$

$$\beta_{s,J} = \sum_{\substack{s_j \in \{0,\ldots,N\},\\ \sum_J s_j = s}} \left( \prod_{j \in J} \beta_{s_j,j} \right), \quad J \subseteq \{1, \ldots, n\} \tag{51}$$

$$\alpha_{s,j,i} = \sum_{\substack{K \subseteq \mathcal{F}_j,\\ i \in K, |K|=s}} \tilde{U}_j(F_K), \quad j = 1, \ldots, n, i \in \mathcal{F}_j \tag{52}$$

$$\alpha_{s,j,i^*} = \sum_{\substack{K \subseteq \mathcal{F}_j,\\ i \notin K, |K|=s}} \tilde{U}_j(F_K), \quad j = 1, \ldots, n, i \in \mathcal{F}_j \tag{53}$$

Definer ligeledes for kvotienterne $a_K$ i Korollar 2,

$$a_s^* = \sum_{x=1}^{N-s} (-1)^x \frac{1}{s+x} \binom{N-s-1}{x-1} \tag{54}$$

$$a_s = \sum_{x=0}^{N-s} (-1)^x \frac{1}{s+x} \binom{N-s}{x} \tag{55}$$

Da kan vi udregne

$$V(F_i) = \sum_{s=0}^{N} \sum_{s_1=0}^{N-s} \beta_{s-s_1,\{1,\ldots,N\}\setminus\{j_i\}} (\alpha_{s_1,j_i,i} a_s + \alpha_{s_2,j_i,i^*} a_s^*) \tag{56}$$