


雇用政策評価へのサバイバル 分析の活用

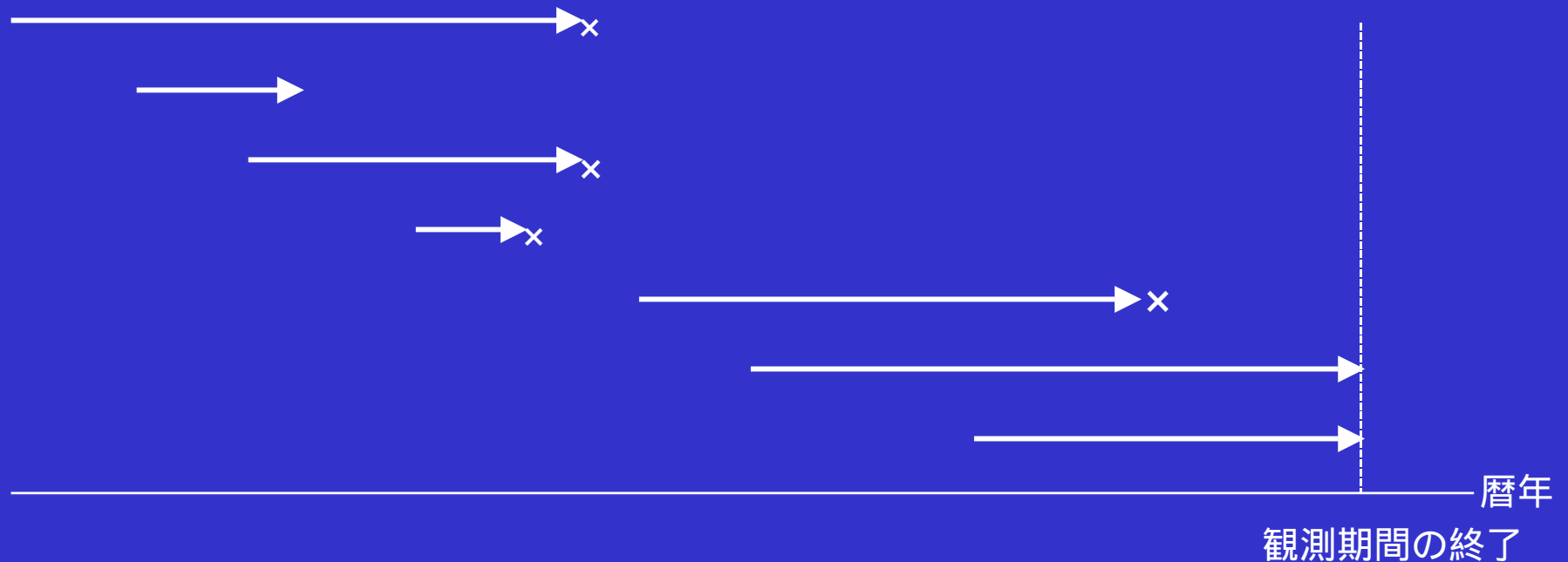
大阪大学社会経済研究所
大日康史

サバイバル分析

- ・不可逆的な状態間変化が対象  期間の分析に適している
- ・最初は医学分野(心臓移植)に適用
- ・観測期間の打ち切りにも頑健
- ・ノンパラメトリック、セミパラメトリック、パラメトリック、split populationモデル、sequential probitタイプとモデルも様々
- ・雇用政策との関連でも失業期間、新規開業からの存続期間、フリーター等の応用例。

Kaplan-Meyer estimator(ノンパラメトリック)

仮想例



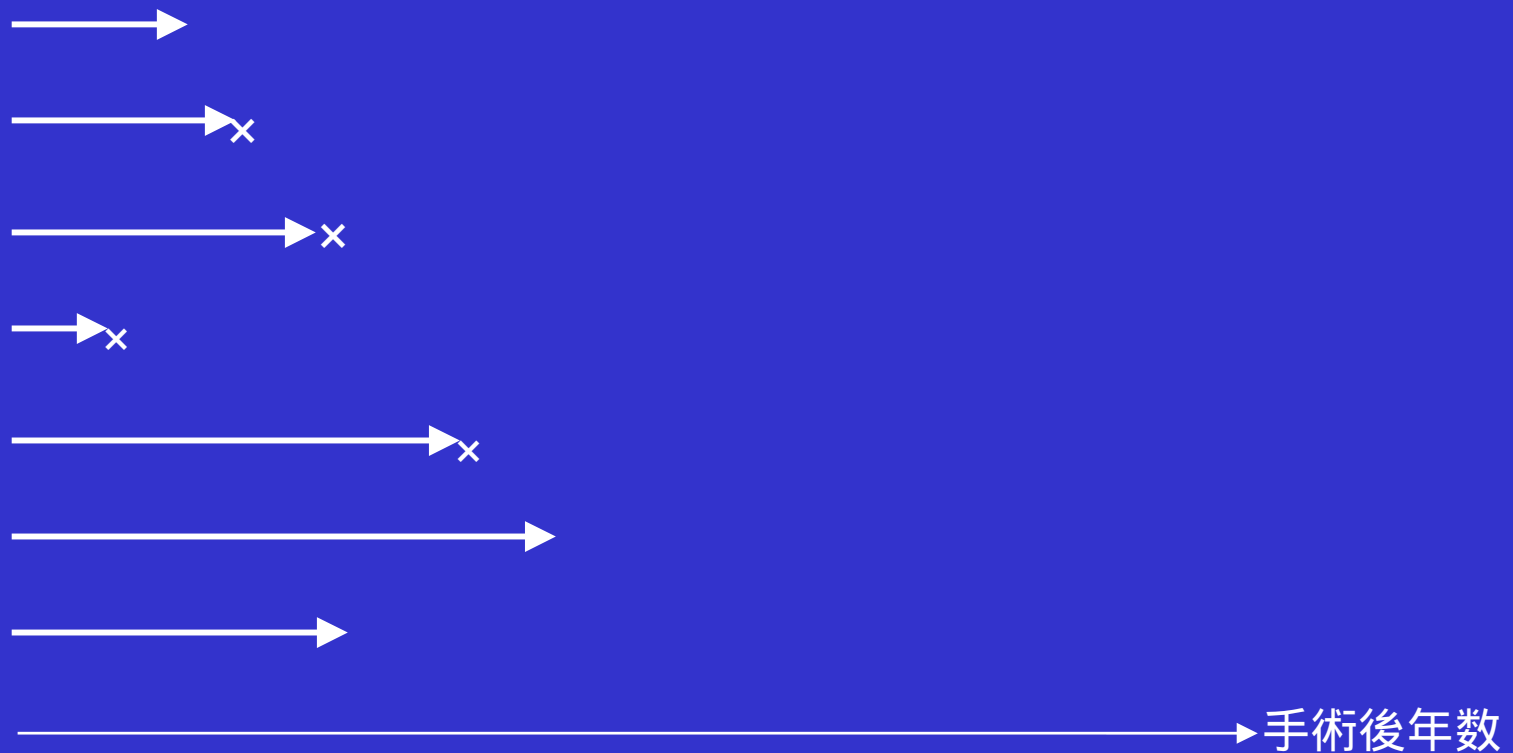
線分の始点: 心臓移植手術

線分の終点: x, 死亡

, 観測打ち切り (観測期間の終了、生存/死亡の未確認)

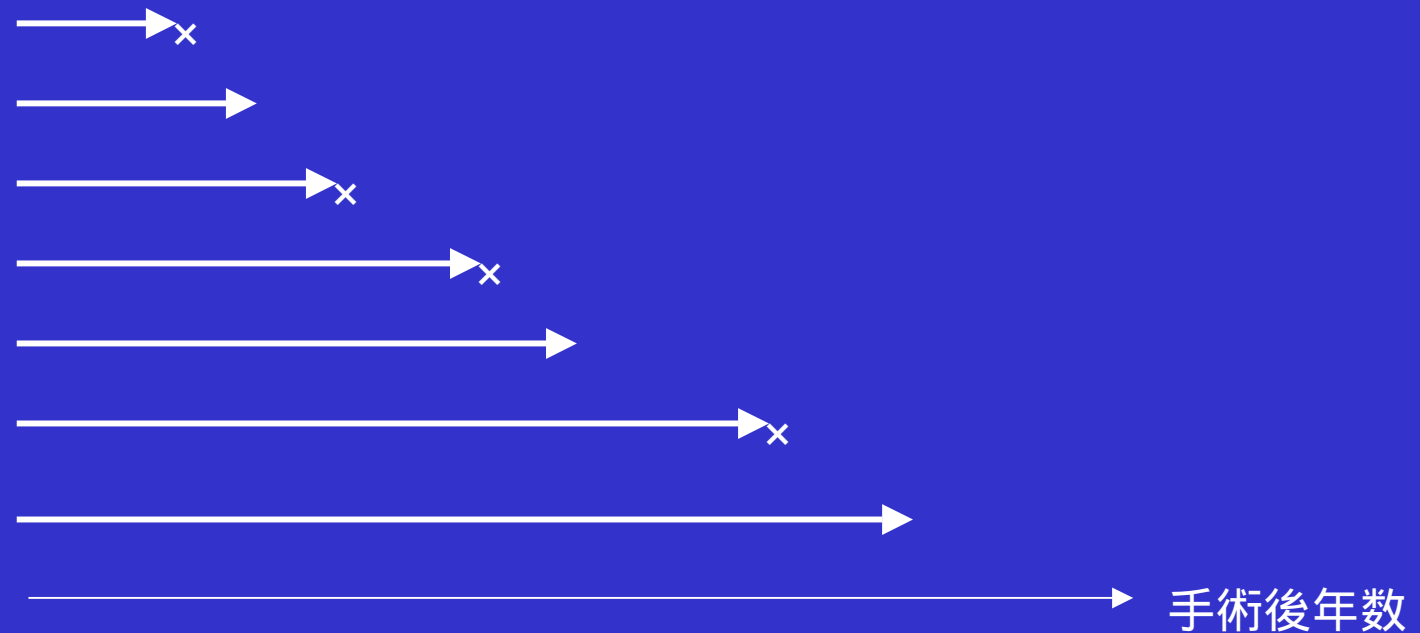
Kaplan-Meyer estimator(つづき)

・期間を手術後年数とする



Kaplan-Meyer estimator(つづき)

・期間の長短で並び替える



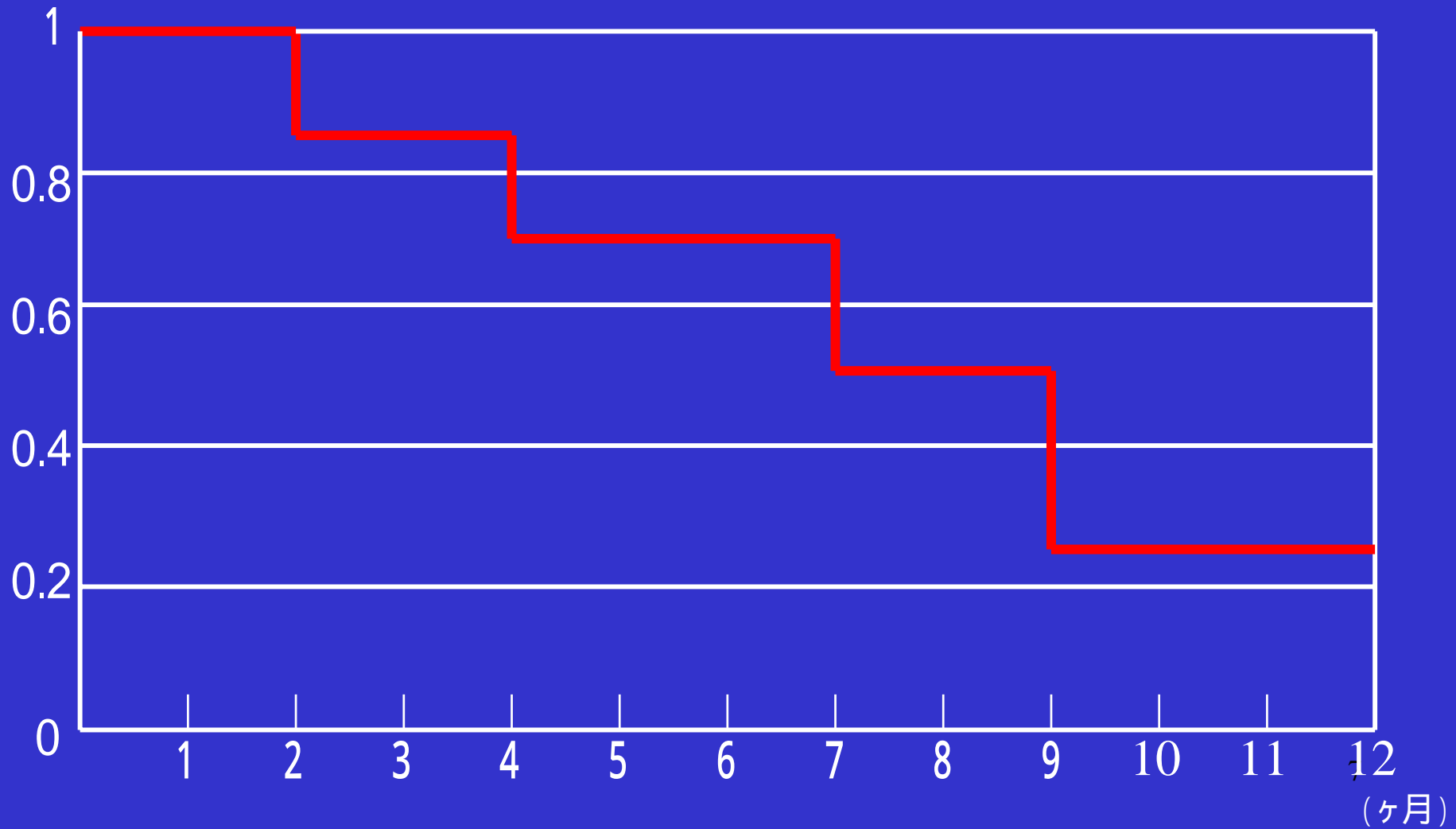
生存確認者数	7	7	7	6	5	4	4	3	2	2	1	1
死亡者数	0	0	1	0	1	0	1	0	0	1	0	0

Kaplan-Meyer estimator(つづき)

生存期間	1	2	3	4	5	6	7	8	9	10	11	12
生存確認者数	7	7	7	6	5	4	4	3	2	2	1	1
死亡者数	0	0	1	0	1	0	1	0	0	1	0	0
死亡確率	0	0	1/7	0	1/5	0	1/4	0	0	1/2	0	0
累積死亡確率	0	0	.14	.14	.31	.31	.49	.49	.49	.75	.75	.75
生存確率	1	1	.86	.86	.69	.69	.51	.51	.51	.25	.25	.25

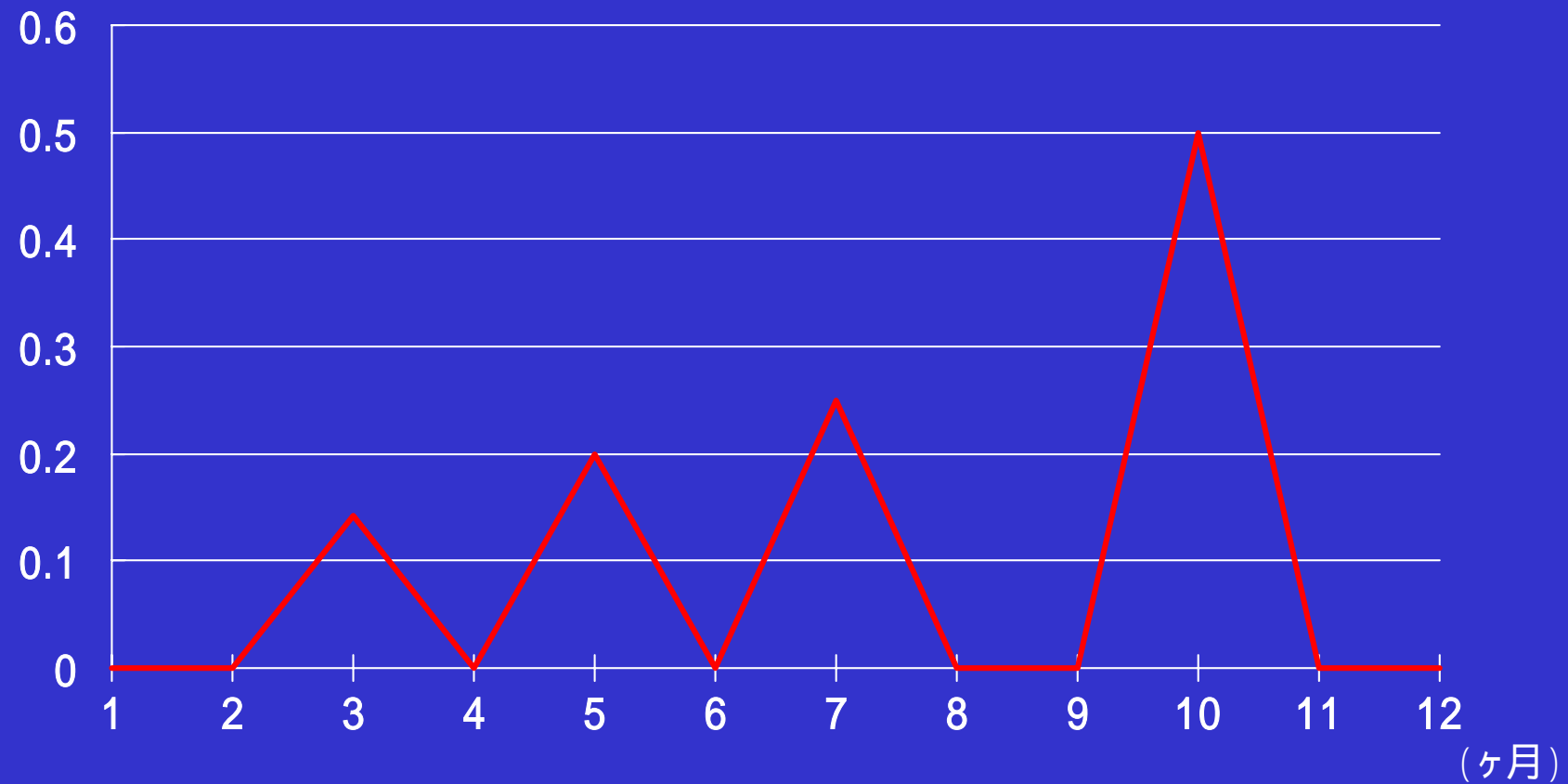
Kaplan-Meyer estimator(つづき)

・サバイバル関数

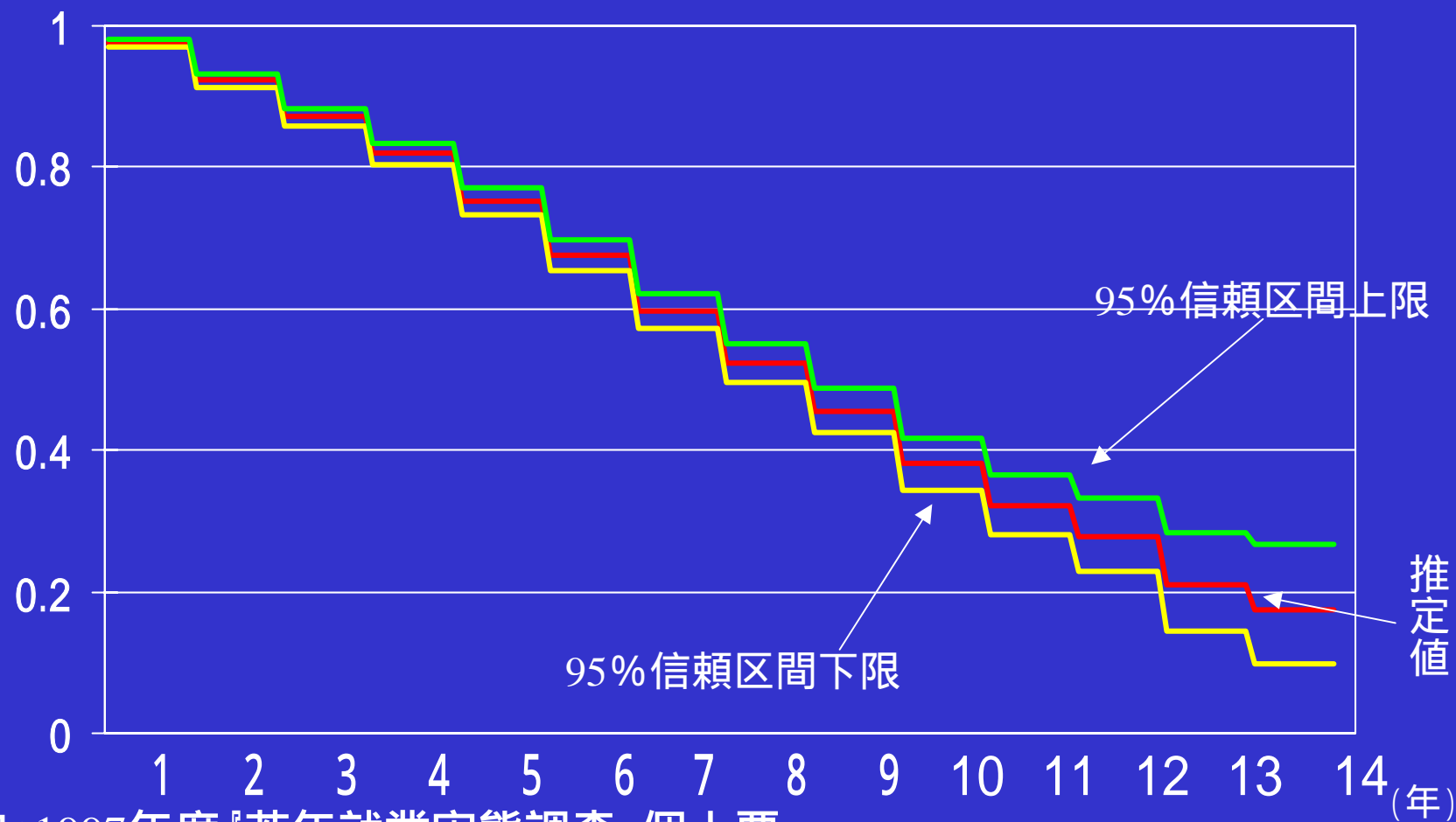


Kaplan-Meyer estimator(つづき)

・ハザード関数



Kaplan-Meyer estimator(具体例): 卒業後正社員就職までの期間



出典: 1997年度『若年就業実態調査』個人票

Kaplan-Meyer estimatorの利点と欠点

- ・利点 サバイバルあるいはハザード関数の関数型を特定化していないので、特定化のミスは排除できる
- ・欠点 特定の要因のサバイバルあるいはハザード関数への影響を評価できない
外挿できないために予測には不適切

Cox比例ハザードモデル(セミパラメトリック)

- ・ハザード関数に比例的に影響を与える要因を推定

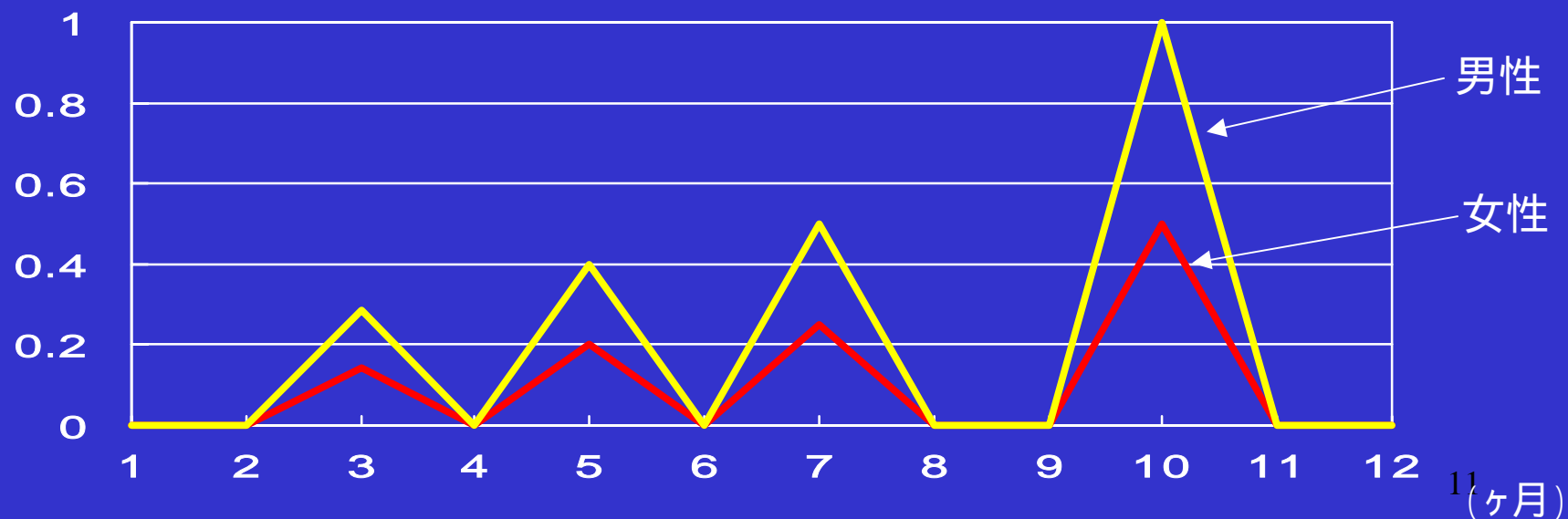
$$h(t) = h_0(t) \text{EXP}[X \beta]$$

$h(t)$:ハザード関数

$h_0(t)$:ベースラインのハザード関数

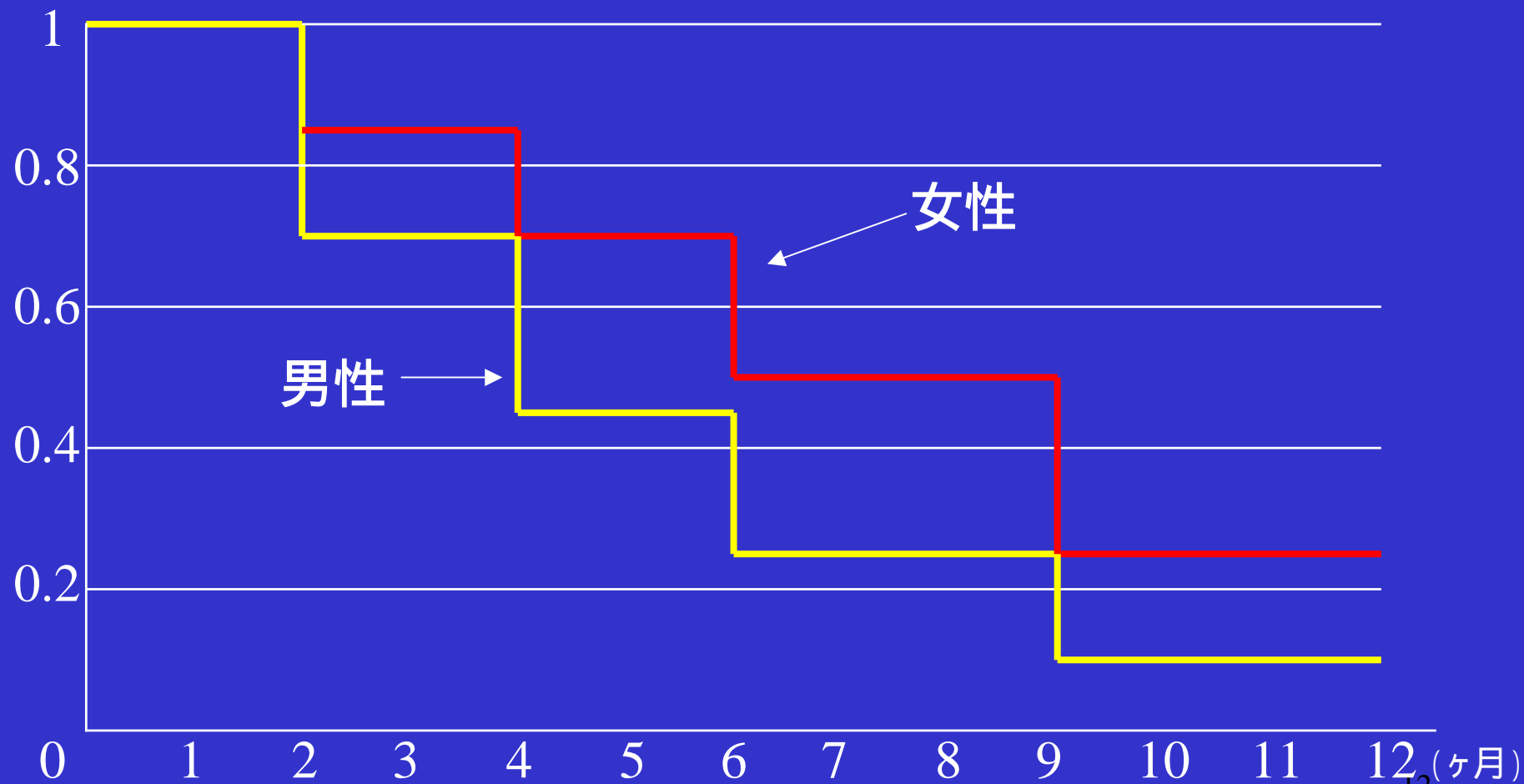
X :ハザード関数に比例的に影響を与える要因

- ・例:男性の方が死亡確率が倍である場合、(X は男性ダミーのみ)



Cox比例ハザードモデル(つづき)

・サバイバル関数



Cox比例ハザードモデル(具体例)

:日本における失業期間の推定

	推定値	確率値
女性ダミー	-.1247524	0.000
年齢	-.0050528	0.623
年齢 ²	.0000855	0.553
失業給付ダミー	-.8645787	0.000
自己都合退職ダミー	-.2979061	0.000
資格ダミー	.0252722	0.321
失業給付終了期ダミー	.2230663	0.000

出典: 拙稿(2000), "An Empirical Evidence of Moral Hazard due to Unemployment Benefits," ISER DP no.511.

Cox比例ハザードモデルの利点と欠点

- ・利点 ベースラインのサバイバルあるいはハザード関数の関数型を特定化していないので、特定化の誤りは排除できる
特定の要因のサバイバルあるいはハザード関数への影響を評価できる
- ・欠点 比例性の仮定が非常に厳しい・・・多くの実証例では満たされない = 影響が時間とともに変わりうる
外挿できないために予測には不適切

パラメトリック推定

- ・サバイバル/ハザード関数を特定化
- ・具体例

	指数分布	ワイブル分布	対数ロジスティック分布
サバイバル関数	$\exp(-t)$	$\exp(-t^\alpha)$	$\frac{1}{1+t^\alpha}$
ハザード関数		$t^{\alpha-1}$	$\frac{t^{\alpha-1}}{1+t^\alpha}$

パラメトリック推定の利点と欠点

- ・利点 比例性の仮定には縛られない
外挿できるために予測には適切
- ・欠点 特定化の誤りの危険大

最近のモデル

- ・Competing Riskモデル: 変化後の状態が複数あるモデル
 - ・・・一方の状態へのハザードを観測打ち切りと見なして状態毎に推定
 - ex. 入院期間の分析における退院と死亡
 - 失業期間の分析における再就職と非労働力化
- ・Left Censoring: ある状態に入った起点が不明な場合
 - ・・・パラメトリックモデルで条件付けしながら推定
- ・Split Populationモデル: 最後までハザードしない標本が多い(重要である)場合に適したモデル
 - ・・・いつかハザードするかどうかを同時に考慮しながら推定する
 - ex. フリーター(卒業後正社員になるまでの期間分析)?

最近のモデル(つづき)

- Sequential Probitモデル: 比例性の仮定や特定化の誤りから免れたモデル

…各期にハザードするか否かをprobit(logit)で表現し、パネルデータとして分析

cf. Gilleskie, D.B. and T.A. Mroz (2000), "Estimating the Effects of Covariates on Health Expenditure," NBER working paper #7942.

具体例: フリーターの分析

	推定値	比例性の仮説検定
女性	1.239792	0.2225
就職指導	3.660109	0.0087
年齢	0.6683568	0.0285
失業率	0.8609732	0.1715
モデル全体での比例性の仮説検定		0.0001

フリーターの分析(つづき)

卒業後年数1	-.31782797***	卒業後年数11	-.48785955
卒業後年数2	-.32718042***	卒業後年数12	-.37826592***
卒業後年数3	-.34514946***	卒業後年数13	-1.5044112
卒業後年数4	-.36183956***	卒業後年数14	-1.3997562
卒業後年数5	-.40162572***	女性ダミー	-.00450129
卒業後年数6	-.39417511***	就職指導ダミー	.22383317***
卒業後年数7	-.41500965***	就職指導・1年後以降ダミー	-.13090053***
卒業後年数8	-.45232925***	年齢	.00439812***
卒業後年数9	-.413519***	失業率	-.02389741***
卒業後年数10	-.46241638		