

以下は、Kaplan-Meier法による生存曲線の作成例を示している。なお、「データ」のシートでは一切編集を行わず、計算等の作業は「作業・加工」のシートで行う。

▶ 手順・流れ

- ① ワークシートのidごとに、ある疾患の発生の有無を示す変数：case (I列) と観察人年を示す変数：pyear (J列) を作成する。caseには、追跡期間中に該当のidで疾患が発生した場合は「1」、そうでない場合には空欄を入力する。Excel関数で処理する場合には、セルI3に「=IF(E3="", "", IF(E3>0, 1, 0))」と入力した後にコピーし、I4からI12までペーストする。pyearには、追跡開始から追跡終了あるいは疾患発生までの期間を算出する。Excel関数で処理する場合には、セルJ3に「=IF(E3="", D3-G3, E3-G3)」と入力した後にコピーし、J4からJ12までペーストする。
- ② 昇順にならべた追跡年数 (L列) ごとに、該当する人数の合計を算出する (M列)。セルM3に「=COUNTIF(\$J\$3:\$J\$12, L3)」と入力した後にコピーし、M4からM12までペーストする。
- ③ 昇順にならべた追跡年数ごとに、該当する人数の累積和を算出する (N列)。セルN3に「=SUM(\$M\$3:M3)」*1と入力した後にコピーし、N4からN12までペーストする。
- ④ 昇順にならべた追跡年数ごとに、caseの合計を算出する (O列)。セルO3に「=SUMIF(\$J\$3:\$J\$12, L3, \$I\$3:\$I\$12)」と入力した後にコピーし、O4からO12までペーストする。
- ⑤ それぞれの追跡年数時点において、疾患の発生するリスクのある人数 (number at risk) を算出する (P列)。セルP3に「=IF(L3=1, SUM(\$M\$3:\$M\$12), SUM(\$M\$3:\$M\$12)-N2)」と入力した後にコピーし、P4からP12までペーストする。
- ⑥ それぞれの追跡年数時点において、疾患の発生数をリスクのある人数で除して、疾患の発生する確率を算出する (Q列)。セルQ3に「=O3/P3」と入力した後にコピーし、Q4からQ12までペーストする。
- ⑦ それぞれの追跡年数時点において、疾患の発生しない確率を算出する (R列)。セルR3に「=(P3-O3)/P3」と入力した後にコピーし、R4からR12までペーストする。なお、疾患の発生する確率 (手順⑥) を用いて、セルR3に「1-Q3」と入力してもよい。
- ⑧ 累積生存確率 (cumulative probability of survival) *2 を算出する (S列)。セルS3に「=IF(L3=1, R3, R3*S2)」と入力した後にコピーし、S4からS12までペーストする。ここで算出した累積生存確率は、「表示」のシートに Kaplan-Meierプロットとして表示される (ワークシート-図1)。
- ⑨ セルT3に「=(1-S3)*100」を入力した後にコピーし、T4からT12までペーストすると、累積罹患率を算出できる (T列)。今回の例題の場合、アウトカムは死亡の発生ではなく疾患の発生であるので、累積生存確率というよりも累積罹患率 (ワークシート-図2) として可視化するほうがわかりやすい。

* 1 関数中の2つのセルの内、1つは絶対参照、もう1つは相対参照になっている点に注意すること。

* 2 生存関数 (survival function) と呼ばれる。なお、死亡と疾患のいずれの新規発生をアウトカムとした場合も計算手順は同様であるが、疾患の場合は1から累積生存確率を引いた値を可視化することが多い (プロットのy軸が反転する)。