3-3-2-2-2-2-2 年齢が発病に影響を与える場合の年齢情報組み込みモデル

条件設定

① Genotype　G1 と G2 が存在する Gi 群の人口を Pi で表す。Pi は t の関数である。

② Genotype Gi 群は疾患 X を発病した群 Gdi と、発病しない群 Ghi に分けられる。それぞれの群の人口は Pki(k=d or h)で表す。

③ Gki(k=d or h)群の死亡率を年齢 t の関数として Dki(t)と表す。

④ 新規発症は非罹患群 Ghi からのみおきるとして、その発症率は年齢の関数として Rhi(t)と表す。ケース vs 非ケースのケース・コントロールスタディを行うとして、それぞれの群から得られる年齢情報 Ta(解析時年齢)、Tb(採血時年齢)、Tc(臨床情報取得時年齢)、Td(発病年齢)

$$Ph_1(t) + Ph_2(t) + Pd_1(t) + Pd_2(t) = 1$$
$$Phi(T) = Phi(o) - \int_0^T Phi(t) \times Dhi(t)dt - \int_0^T Phi(t) \times Rhi(t)dt$$
$$Pdi(T) = Pdi(o) - \int_0^T Pdi(t) \times Ddi(t)dt + \int_0^T Phi(t) \times Rhi(t)dt$$

ある年齢Tにおける人口がTの関数で与えられたケースを、このような集団からサンプリングすると、Genotype I の占める率

$$Fdi(t) = \frac{Pdi(t)}{Pd_1(t) + Pd_2(t)}$$

コントロールをこのような集団からサンプリングすると、

$$Fhi(t) = \frac{Phi(t)}{Ph_1(t) + Ph_2(t)}$$

但し、このときの T は、臨床情報取得時年齢であるので、Tc を用いる。

以上が一般的なケース・非ケース集団からのサンプリングを行った場合である。

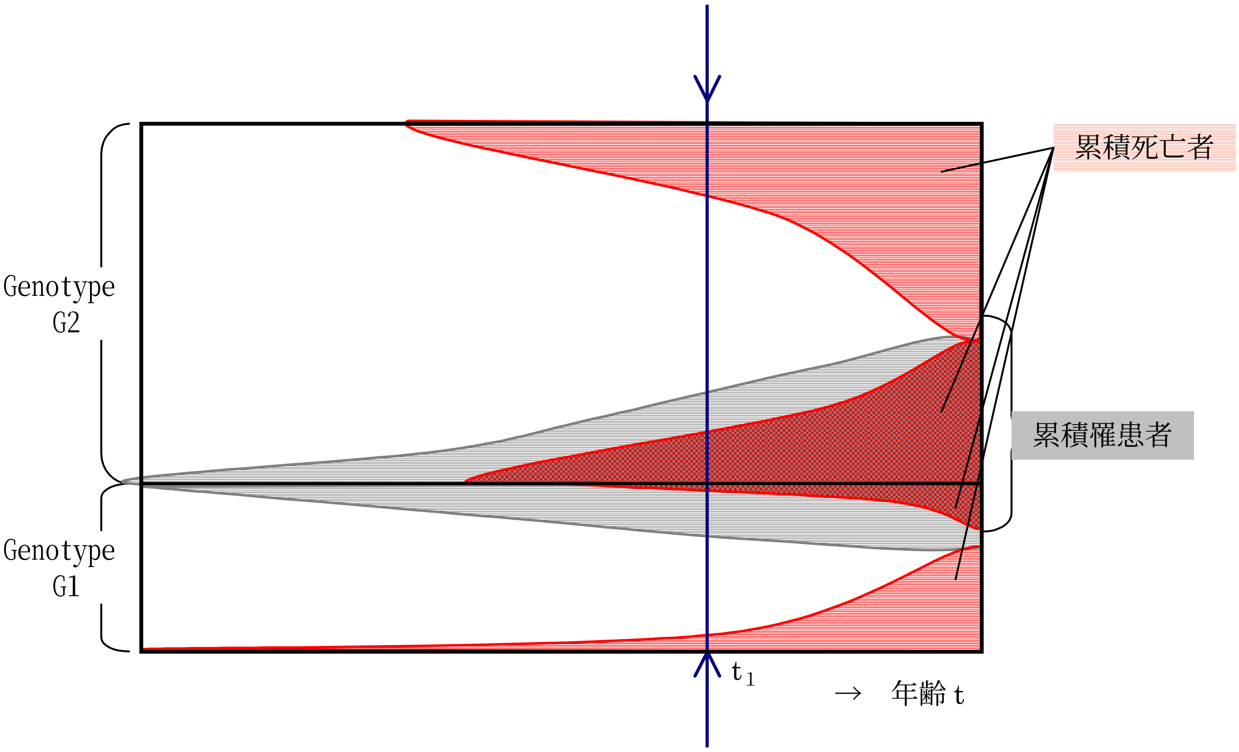
上記のモデルについて、実際のデータを解析することを考える。

ケースデータが $Tc=t_j$ につき Ndi 人、非ケースのデータが同様に Nhi 人　観測されたとすると、このような観測データが得られる。

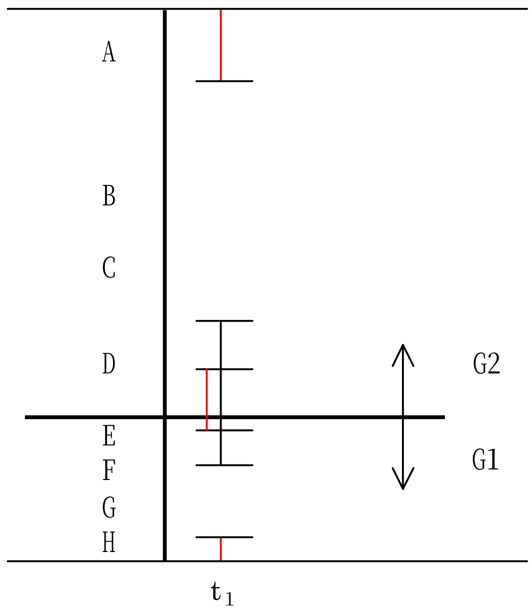
対数尤度 L は

$$L = \sum_T Nd_1(T) \ln Fd_1(T) + \sum_T Nd_2(T) \ln Fd_2(T) + \sum_T Nh_1(T) \ln Fh_1(T) + \sum_T Nh_2(T) \ln Fh_2(T)$$

ある観測データを基に最尤推定を行い、T の関数 $Rh_1(T)$と $Rh_2(T)$とに有意な差が認められれば、Genotype1/2 は発症に関連していると検定される。



年齢別 t_1 でのサンプリング



A：Genotype G2 で年齢 t_1 に到るまで、ある疾患に罹患せず死亡した画分

B：Genotype G2 で年齢 t_1 にて、ある疾患に罹患せず存在している画分

C：Genotype G2 で年齢 t_1 までに、ある疾患に罹患し生存している画分

D：Genotype G2 で年齢 t_1 までに、ある疾患に罹患し、かつ死亡している画分

E：Genotype G1 にて G2 の D に相当する画分

F：Genotype G1 にて G2 の C に相当する画分

G：Genotype G1 にて G2 の B に相当する画分

H：Genotype G1 にて G2 の A に相当する画分

簡略化したモデルを用いてみる。

Genotype1/2、疾患の有無 d/h は死亡率に無関係であるり、60 歳まで死亡率は 0、以後加齢比例して死亡率は上昇する。

$Dd_1(t) = Dd_2(t) = Dh_1(t) = Dh_2(t)$

$$\begin{cases} 0 \leqq t \leqq 60 & D**(t) = 0 \\ t < 60 & D**(t) = d(t-60) \end{cases}$$

また、20 歳以前の発病はなし、以後加齢に比例して発症率は上昇し、Genotype G2 の発症率は G1 のそれの r 倍と固定されているとする。

$Rh_2(t) = cRh_1(t)$

$Rh_1(t) = r(t-20)$　　$(0 \leqq c \leqq 1)$

$Pd_1(0) = 0$

$Pd_2(0) = 0$

$Ph_1(0) = P$

$Ph_2(0) = 1 - P$

簡略化モデルの概略図