Unobserved Heterogeneity

The life table summarizes the experience of a population. This is not representative of the experience of individuals unless these are homogenous. This unit deals with the consequences of unobserved heterogeneity, with an application to the mortality cross-over in Black and White mortality in the U.S. A classic reference is Vaupel, Manton and Stallard (1979).

The Multiplicative Frailty Model

A popular approach to modeling unobserved heterogeneity assumes that the hazard $\mu_i(x)$ for individual i at age x is the product of two terms, a baseline hazard $\mu_0(x)$ and a multiplicative term θ_i representing the individual's frailty, so

$$\mu_i(x) = \mu_0(x)\theta_i$$

A person with $\theta = 1$ represents the baseline risk. A person with $\theta = 1.5$ has 50% *higher* risk than our reference individual *at every age.* A person with $\theta = 0.5$ has 50% *lower* risk than the reference individual. The formulation is just like a proportional hazards model, except that we don't observe a person's frailty.

Let $p_i(x)$ denote the probability that individual i will survive to age x,

$$p_i(x) = \Pr\{X > x\} = \frac{l_i(x)}{l_i(0)}$$

This is just $l_i(x)$ if the radix is 1. From our results relating survival probabilities to hazards we have

$$p_i(x) = p_0(x)^{\theta_i}$$

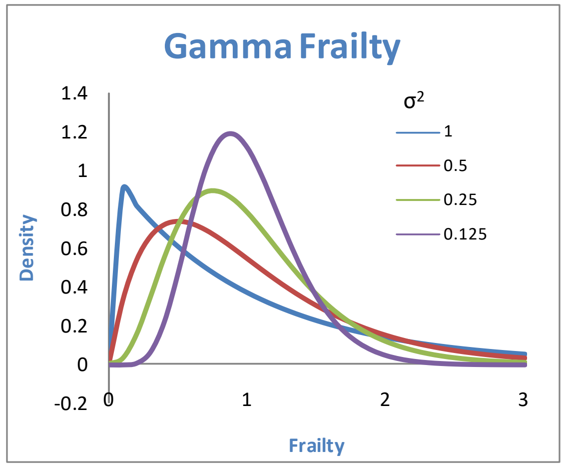
where $p_0(x)$ is the baseline survival probability. This follows from writing the survival probability as $p_i(x) = \exp\{-\int_0^x \mu_i(a)da\}$ and then substituting the model for the individual hazard. So if the reference individual has an 80% change of living to age 60, one with $\theta = 1.5$ has only a 72% chance, whereas one with $\theta = 0.5$ has an 89% chance.

Gamma Frailty

The next step is to assume that frailty has a distribution in the population. A common assumption is to postulate a gamma distribution, which has density

$$g(\theta) = \frac{\theta^{\alpha-1}e^{-\beta\theta}\beta^{\alpha}}{\Gamma(\alpha)}$$

with parameters α and β . The mean is α/β and the variance is α/β^2, so we often set $\alpha = \beta = 1/\sigma^2$ to get a mean of one and a desired variance. The figure on the side shows gamma densities with mean one and variances of one, a half, a quarter and 1/8.



One could use other distributions for frailty, and results have been obtained for discrete mixtures and for cases where frailty has an inverse Gaussian or a compound Poisson distribution, but gamma is by far the most popular choice.

Population Average Survival

The survival function we estimate with a life table is an average for individuals with different frailties. Suppose the entire population consisted of just the three individuals in the initial example. Then the average probability of living to age 60 would be 80.3%, the average of 80, 72 and 89.

Of course a population will have more than three individuals, so we average using the distribution of frailty. The average survival probability in the population is then

$$p(x) = \int_0^\infty p_0(a)^\theta g(\theta) d\theta$$

In general this is not the same as the baseline. We call $p(x)$ the *population-average* survival.

If frailty has a gamma distribution with mean one and variance σ^2, then with a bit of algebra one can show that the population survival is given by

$$p(x) = \frac{1}{[1 + \sigma^2 H_0(x)]^{\frac{1}{\sigma^2}}}$$

where $H_0(x) = \int_0^x \mu_0(a) da$ is the integrated baseline hazard. This is a Pareto distribution of the second kind.

Survival functions are useful but not terribly informative, so I turn attention to the hazard.

Population Average Hazard

To compute the population hazard we proceed from first principles, taking the negative log of the survival probability to obtain a cumulative (or integrated) hazard and then differentiating to obtain the hazard. If we follow that approach it can be shown that

$$\mu(x) = \mu_0(x) E(\theta | X > x)$$

where $E(\theta | X > x)$ is the expected value of frailty among survivors to age x.

A more specific result can be obtained if we assume that frailty at birth has a gamma distribution with mean one and variance σ^2. In that case the mean frailty of survivors to age x is

$$E(\theta|X > x) = \frac{1}{1 + \sigma^2 H_0(x)}$$

where $H_0(x)$ is the integrated baseline hazard, as before. Thus, under gamma frailty the population average hazard is

$$\mu(x) = \frac{\mu_0(x)}{1 + \sigma^2 H_0(x)}$$

At birth mean frailty is one and the population average hazard is the same as the baseline individual hazard. As time goes by, however, the mean frailty of survivors declines, becoming less than one, and as a result the population average hazard becomes lower than the baseline individual hazard. This can be seen from the fact that the integrated hazard, which increases with age, is in the denominator of the formulas for the mean frailty and the population average hazard.

Our interpretation is this result is that the frail tend to die first, so over time the population becomes increasingly selected, consisting of individuals who are more robust. Note also that frailty declines faster (so selection operates more quickly) when the population is more heterogeneous to start with (larger σ^2) or the risk is higher (larger baseline hazard $\mu_0(x)$ and hence larger $H_0(x)$).

Example: To fix ideas consider a situation where the hazard is constant over time for each individual but there is heterogeneity of frailty. Specifically suppose the individual hazard is $\mu_0\theta$, where μ_0 is the baseline hazard and θ denotes frailty. If frailty has a gamma distribution then the population hazard is

$$\mu(x) = \frac{\mu_0}{1 + \sigma^2 \mu_0 x}$$

and declines from μ_0 at birth to zero as $x \to \infty$. It will decline faster for larger μ_0 or larger σ^2.

A constant hazard model doesn't work well for mortality but it approximates other situations, such as time to conception among fecund women trying to conceive a child. Assume that each woman's fecundability is constant over time, at least for a few months, but women differ in their fecundability. According to these results the population hazard would decline over time even though it's constant for each woman. This occurs because more fecund women tend to conceive first, and the survivors become increasingly selected for lower fecundability.

When frailty at birth has a gamma distribution one can show that the distribution of frailty among survivors to age x is also gamma with the mean given above and variance

$$var(\theta|X > x) = \frac{\sigma^2}{\left(1 + \sigma^2 H_0(x)\right)^2}$$

Note that the variance at birth is σ^2 but over time the variance get smaller and smaller, so the population becomes more homogeneous. Frailty, of course, also declines. An interesting feature of gamma frailty is that the coefficient of variation (standard deviation over mean) remains constant.

The Inversion Formula

So far we have gone from individual to population hazards. Can we go the other way? The answer is yes, and leads to interesting applications.

If frailty has a gamma distribution then one can show that the baseline hazard satisfies

$$\mu_0(t) = \mu(x)e^{\sigma^2 H(x)}$$

where $H(x) = \int_0^x \mu(a)da$ is the cumulative (integrated) population hazard. (The negative log of the population survival function with radix 1.)

Example. We considered earlier how a gamma mixture of exponentials leads to a declining population hazard. I now show that a constant population hazard can be viewed as a mixture of something else. If the population hazard is constant then $\mu(x) = \mu$ and the cumulative hazard is $H(x) = \mu x$. Plugging these functions into the inversion formula we find that the baseline individual hazard is

$$\mu_0(x) = \mu e^{\sigma^2 \mu x}$$

an exponential function of x which we recognize as a Gompertz hazard. Thus, we have the remarkable result that a population that shows a constant hazard may result from individuals with gamma distributed heterogeneity who face hazards that increase exponentially with time.

The Identification Problem

You may begin to suspect that we have a bit of an identification problem here, because a flat population hazard could also result from a homogeneous population where each individual's hazard is flat. All we can estimate is hazards for populations, or groups. It pays to be aware, however, that the hazards for the individuals may be different. In particular, we can't distinguish heterogeneity from negative duration dependence.

The Mortality Cross-Over

The online computing logs illustrate these ideas with an application to U.S. mortality. We start with population average survival and hazard curves for blacks and whites and note the well-documented mortality cross-over. We then use the inversion formula to find subject-specific hazards for blacks and whites that do not cross, yet under heterogeneity lead to population-average hazards that do cross. The underlying explanation is that blacks face higher mortality at younger ages and hence become more highly selected at older ages. The alternative explanation is age misreporting, which may be particularly prominent among older blacks because of the lack of birth certificates.