Session 5B: Mortality Measurement Q&A

Presenters: Natalia S. Gavrilova
Bob Howard
Paul J. Sweeting

Presented at the Living to 100 Symposium
Orlando, Fla.

January 5-7, 2011

Kai Kaufhold: First of all thanks to Natalia for the compliment that actuaries might have gotten something right once. It might be a coincidence because might it have been that Gompertz was arbitrarily used as the extrapolation method? Or, could you tell us what the extrapolation method was that got the 1900 cohort life table right?

Natalia Gavrilova: They used methods in which they actually extrapolated qx not mortality force. In using the method, they assumed that after age 95 probabilities of death increased every year of age by 5 percent for men and 6 percent for women. I was surprised that it coincided with real data because they used extrapolation and did not fit real data. I was surprised by this coincidence. But I would just like to emphasize that if you would like to model mortality at very old ages, you probably should not use probability of death because mortality force follows Gompertz model very well, at least according to our data. What was actually most interesting to me from our work was that mortality force and probability of death diverge significantly after age 100 so they are not the same. You need to have different models for them.

Kai Kaufhold: Thank you. It actually goes to show that a simple method used by a practitioner might actually hit as well and that's why I also thought Bob's presentation was very interesting.

I have a question for Paul. Essentially what you're trying to do is reduce the importance of your rating factor QXs or essentially getting rid of the betas in your credibility formula as much as you can. Couldn't you increase the bandwidth by using a stronger assumption than the simple log linear approximation? For instance, a Gompertz relationship like Natalia is suggesting for older ages?

Paul Sweeting: You could and it doesn't really matter what form you believe mortality rates follow as long as it's fairly simple. The danger is if your model is too complex, then you're fitting it too much to the data points. If you're looking at age 65 and say that the mortality rate for age 65, the number of deaths for age 65 is actually a little on the low side, but the number of deaths at age 64 and 66 were a little on the high side. Well, if you fit a line to those mortality rates which follows that a little too much, then you actually lose that information. What you want to do is something which captures the broad shape of mortality and reflects the fact that while you observe mortality at the age of interest is a little low. Once you take into account the information from the ages on the side, you get to a more accurate mortality rate. But you could use any law, provided it's a reasonably simple law linking the mortality rates. The log linear approach is similar to the Gompertz approach anyway so this is a long explanation. It's the same sort of thing.

Kai Kaufhold: But aren't you essentially replacing one prior assumption with another one? I'm turning around 180 degrees, but when you're using an experience-rated mortality that you get from some postcode model, then essentially you're putting a prior assumption into it and you're replacing that with another prior assumption that locally your mortality by age is log linear or whichever.

A: You're actually right because it's not some sort of magic bullet which means your data are automatically better, but one of the issues that I've seen with things like postcode rating is that they can't capture all of the information and some of them don't even capture the majority of the information. There's a lot of variability in the mortality rates, which just aren't picked up by postcode. A good example might be if you're looking at mortality rates for a pension plan for people working in a particular factory, in a particular industry. Now, where their working is going to have a major impact on their mortality rate and if it so happens that you got a well paid sheet metal worker who is living on the same street as a bunch of accountants, then the postcode, for example isn't going to pick up that information and those plan-specific details are more likely to be picked up if you're looking at the mortality rates over the whole mortality curve rather than just at the rate. So whilst I say the beta approach is nice because it's nice and objective, so there is subjectivity to the fact you're using this extension approach at all about the bandwidth, around what your external risk measure is so there's still a whole bunch of things you need to decide. It's better than just looking at the number of deaths at this particular age and then a whole lot of external stuff.

Kai Kaufhold: Great. Thank you very much.

Tom Edwalds: I have a combination of comments and questions. First, it's remarkable that, after almost 200 years, Benjamin Gompertz is still right. That just amazes me. In fact, I remember nine years ago it was Ward Kingkade presenting data from Medicare that was indicating the continued increase in mortality. He was the only one at this conference who was saying there's no mortality deceleration. Now we're getting back to that opinion that mortality deceleration isn't really a true phenomenon. I like Natalia's analysis getting down to the month of birth and being able to replicate the Gompertz curve up to a fairly high age.

A thought I had about that, and this is relevant to both Bob and Natalia, is that, in the United States anyway, the birth cohort starting at about 1900 would have to have provided proof of age when they registered for Medicare. That's almost universal, so we'll actually start getting better data from birth cohort 1900 and on. I guess if you're using the Social Security Death Master File and you're saying you've got the wrong age, since you have an actual date of death and actual date of birth, it must be the date of birth that's wrong. So you're saying we have a misreported date of birth, and that could happen if they never had to prove their actual date of birth and whatever they reported was accepted later on. So I think going forward we're going to get much better data for those high ages because of the requirement in the United States of having the birth date verified when applying for Medicare. If you can comment on that I would appreciate it. I actually have a question for Paul as well.

One of the very common techniques I've seen in actuarial practice for trying to smooth out data is to compare the experience by quinquennial age group to a reference table by calculating A/E ratios, and then interpolate between them. How different is your method from just taking a quinquennial age group and doing that kind of smoothing?

Paul Sweeting: There are a couple of differences. The first is if you just take a quinquennial age group because the number of deaths would almost have to increase exponentially, you have a risk of biasing the mortality rate upwards if you just take that group and you're seeing the midpoints of the group. The other difference is even if you didn't have that bias you're giving equal weight to all ages in that group whereas my approach is saying, well, I know the number of deaths at the age I'm looking at is right and relevant. I have slightly less confidence in the ones on either side, so you're reducing the weightings a bit, but it's not a million miles away. It could be regarded as being a more sophisticated approach or an enhancement to that approach.

Bob Howard: If I were doing as Tom suggested, I'd be using the actual exposure for each individual age and multiplying that by the standard mortality rate for each age and then dividing to get the average, if I needed an average age, I'd be getting the average age that I was working with, so I wouldn't have that bias of a different slope, and if you made that kind of adjustment then how close is it to your method?

Paul Sweeting: Let me think. I still think you would have some lack of linearity there, but I would need to sit down with a pen and paper and figure out how much difference it was. What I really need to do which I wasn't able to do with the work so far is actually apply it to real data and see what difference it makes because the only way to test it is by actually applying it to a real data set and seeing how much the results differ and whether there's any improvements to it.

Bob Howard: Or you could use simulated like I did and then you get the whole distribution.

Natalia Gavrilova: I simply would like to again emphasize, you get Gompertz only if you project and model mortality force and not if you model probability of death. For probability of death you do not get Gompertz.

Tom Edwalds: A question for Bob. In using the method of extinct generations, how different is your method from the way others have done it? I thought there was a standard technique for taking a not quite extinct generation and going backwards and I haven't used it myself so I wasn't sure how different what you have done is from that.

Bob Howard: Well, certainly the technique for extinguished cohorts, that's been around for a century at least so there's nothing new there. For the non-extinguished cohorts, it could well be just my ignorance and I've not run across something. As I said, I'm not one of the academics up here so it's not part of my job to go out and read all the stuff that's out there and I'm years behind catching up. So it could be that somebody has already published something that's doing the same sort of thing but better than what I did. What I put forward is something because I wasn't aware of anything else and it seemed to work and that's why I presented it.

Tom Edwalds: Thank you.

Jordi Posthumus: Just a quick question for Natalia. In trying to identify an ultimate Q, would there not be some sort of value to the following approach. If you took the data that you guys have and you said let's look at everybody who died between ages say 60 and 70 and you entirely exclude everybody who did not die during that time period and you calculate an ultimate *Q* for that group so having no exposure for people who died in any other age groups and then you do the same thing for people who died from age 70 to 80 and the same thing for people who died from age 80 to 90 so you have completely separate cohorts and you look at what the ultimate *Q* in these groups is and the trajectory of these sets, maybe you could infer something about *Q*s in the extreme old ages where you have very low data. I mean I have no clue what such a *Q* would be. Would it be just random or would it converge or is it something you've looked at or not looked at or just a question?

Natalia Gavrilova: What is the ultimate *Q*? I don't know some of the actuarial terminology.

Jordi Posthumus: Sorry, so if you look at the ultimate realized mortality rate of the group that you know did die between age 90 and 100, and you exclude other groups. Because when you get to age 110 to 120 you have a finite set of folks, right? It's a really small group of people and basically you're trying to estimate the ultimate *Q* by just really using those folk because they're the only folk in the exposure and the only folk in the numerator and if you say, maybe for people who died who had the modal death centered on age 85, so guys who died between age 80 and age 90, that group themselves would have a characteristic ultimate *Q* and so you'd say, okay, you've got people whose modal age of death shifts along. These are distinct populations that each have an ultimate *Q* and perhaps that thing has a trajectory of its own that could be sort of tractable and informative when you get to the very old ages and *Q* being the mortality rate.

Natalia Gavrilova: At least looking at the mortality experience from age 40 to age 60 or so is informative, because the slope of mortality growth estimated at younger ages is the same as the slope estimated by the Gompertz model after age 100. Actually, we did not see that there is so-called two-stage Gompertz.

Leonid Gavrilov: Well, the thing is you cannot do this with ultimate *Q* at young age because there are so many people who are exposed to death at this age. This idea of the method of extinct generations needs to have all people under the exposure. But I can translate your question in another way. So the question is that, of course, in the case of data truncation at age 115 some people may be alive after this age that were not counted because of truncation and for this reason the denominator is underestimated and the mortality rate is overestimated. All this story with Gompertz and all the things you think may be observed because we did not take into account the very tail of survival curve and the answer to this question is that we played with different specifications and initially we applied the method of extinct generations, which included all the garbage that was after age 115 and still there was a Gompertz-like curve. But then we found that there are really some weird claims, some people died at 125 years, so we decided that it is prudent to decrease the garbage, but you are quite right that there might be some small bias resulting in overstating mortality rate and maybe overstating the slope with age caused by this cut in the tail.

Natalia Gavrilova: Actually we did not cut the data because for mortality analysis we estimated survival curve for the whole range of ages, but then fitted mortality in the age interval 88-106 years because mortality doesn't depend on the rest of the curve. It's not a real truncation. This is simply mortality estimation within specific age interval.

Bob Howard: I just wanted to follow up on that. The problem really is that you don't have a truncated interval. You've got an open interval, that's what makes it difficult. As you described, Jordi, the truncated interval, every time you take that truncated interval you have to end up with *Q*1 at the end because everybody will be dead. So I'm not sure you can learn anything knowing that you've got an artificially truncated interval as opposed to one that's actually open.