MTH6157

Example question on statistical tests

An insurance company has been using the same standard table to calculate premium rates and reserves for term insurance policies since 2017. An actuary wishes to examine whether the continued use of the table is still appropriate and uses data for the last year which is summarised in the table below where the standard table forces of mortality are denoted by μ_x^s.

Age, *x*	Exposed to Risk	Actual Deaths	μ_x^s
65	1933	21	0.011
66	1994	20	0.014
67	2143	36	0.018
68	1713	39	0.023
69	2297	56	0.028
70	2073	65	0.034
71	1956	71	0.041
72	1892	91	0.047
73	1898	99	0.055
74	2175	137	0.064
75	2285	163	0.075

(a) Complete a test of the overall goodness of fit of the observed mortality experience to the standard table rates stating clearly the null hypothesis being tested and your conclusion.

We complete the Chi squared test for the overall goodness of fit

H0: the standard table represents the true underlying mortality for these insurance policies

Our test statistic is $X = \sum_{all\ ages} z_x^2$ where $z_x = \frac{actual\ deaths - expected\ deaths}{\sqrt{expected\ deaths}}$

The calculation of the test statistic is set out in the table below,

Age, x	Exposed to Risk	Actual Deaths	mu_x(s)	Expected deaths	z_x	z_x^2
65	1933	21	0.011	21.263	-0.05704	0.003253
66	1994	20	0.014	27.916	-1.49823	2.2447
67	2143	36	0.018	38.574	-0.41444	0.17176
68	1713	39	0.023	39.399	-0.06357	0.004041
69	2297	56	0.028	64.316	-1.03694	1.075251
70	2073	65	0.034	70.482	-0.65298	0.426383
71	1956	71	0.041	80.196	-1.02689	1.054497
72	1892	91	0.047	88.924	0.22015	0.048466
73	1898	99	0.055	104.39	-0.52754	0.278303
74	2175	137	0.064	139.2	-0.18647	0.03477
75	2285	163	0.075	171.375	-0.63975	0.409282
						5.750706

Under H0 the test statistic follows a χ^2 distribution where,

Degrees of freedom = number of ages (as testing a standard table) = 11

$$\chi^2{}_{0.95,11} = 19.675 > 5.751$$

Therefore we do not reject H0.

Note that in the exam you will be given the Chi squared critical values needed.

Overall we conclude from this test that the standard table does represent the underlying mortality experience.

(b) What aspects of goodness of fit might not be revealed by the test in (a) above?

- A large deviation offset by a number of small deviations
- An outlier in the data
- A small overall positive or negative bias
- A lack of independence between the deviations

Note that given the small age range, runs or clumps of deviations of the same sign are not really an issue here.

The test will also not detect any errors in the data [but that is usually taken as read].

(c) Should the actuary be concerned about outliers in the data?

There are two ways to consider this: formally by test and informally by examining the z_x's

Formally we use the standardised deviations test, grouping the z_x's across their range and comparing with the distribution we would expect under $z_x \sim N(0,1)$.

H0: the standard table represents the true underlying mortality for these insurance policies.

Our calculations are,

Range	< -2	(-2,-1)	(-1,0)	(0,1)	(1,2)	> 2		
Observed	0	3	7	1	0	0		
Expected	2%	14%	34%	34%	14%	2%		
Expected	0.22	1.54	3.74	3.74	1.54	0.22		
(O-E)^2/E	0.22	1.384	2.842	2.007	1.54	0.22		8.21314

The test statistic = 8.213

Under H0 this follows a χ^2 distribution on 6 – 1 = 5 degrees of freedom

$$\chi^2{}_{0.95,11} = 11.071 > 8.213$$

Therefore we do not reject H0

According to the standardised deviations test there is not statistically significant evidence for outliers.

Informally we notice no z_x values >2 or <-2 which again suggests no outliers.

(d) Is there evidence of statistically significant under or over estimation of mortality in the continued use of the standard table?

There are 10 negative and 1 positive standardised deviations in the 11 ages.

A negative z_x means actual < expected deaths therefore our concern would be that the standard table now overestimates mortality.

We test this with the signs test.

H0: the standard table represents the true underlying mortality for these insurance policies.

Under H0 we expect ½ of the deviations to be positive.

We seek the probability of what we observe (1/11) or more extreme outcomes

$$\Pr(0\ positive\ out\ of\ 11) = \binom{11}{0}\frac{1}{2}^{11} = 1(0.000488) = 0.000488$$

$$\Pr(1\ positive\ out\ of\ 11) = \binom{11}{1}\frac{1}{2}^{11} = 11(0.000488) = 0.005371$$

Therefore probability of what we observe or more extreme = 0.000488 + 0.005371 = 0.0059

Which is less than 2.5% or 0.025 (for a two tailed test at 95% significance)

Therefore we reject H0.

There is evidence that the standard table overestimates mortality.

Out of interest

$$\Pr(2\ positive\ out\ of\ 11) = \binom{11}{2}\frac{1}{2}^{11} = 55(0.000488) = 0.026855$$

Therefore Pr(2 or fewer out of 11) = 0.000488 + 0.005371 + 0.026855 = 0.0327 > 0.025

So an observation of 2 out of 11 positive z_x's would not have failed the signs test at 95%

(e) What might be the cause of your finding in (d) above?

This is a question about selection. Think about how selection might cause experienced mortality rates to fall over time so that a standard table starts to overestimate mortality. This could be:

- Time selection [the most likely here – rates generally do fall over time]
- Class selection [different circumstances of policyholders in particular if the average policyholder has become wealthier]
- Temporary initial selection [this would be a change in underwriting practices so that only healthier people were admitted for insurance]
- Spurious selection [the effect is just random variation]

Time selection due to medical advances is almost certainly the most likely.

(f) What are the likely financial consequences of this for the insurance company?

These are term insurance policies so from a financial management perspective the insurance company would rather overestimate mortality in premiums and reserves. Therefore the standard table might still be suitable. If the overestimation becomes too large though premiums will not be competitive with other companies using more up-to-date mortality assumptions. Therefore these should be kept under review probably at least annually.