Life Table Relationships

A number of questions ask us to find and manipulate values from life tables. Understanding these relationships is vital

l_x	Lives alive at age x
$d_x = l_x - l_{x+1}$	Deaths from age x to $x+1$
$\mu_x = -\frac{1}{l_x} \frac{d}{dx} l_x$	Rate of death at age x.
${}_t p_x = \frac{l_{x+t}}{l_x}$	Probability of surviving t years
${}_t q_x = \frac{l_x - l_{x+t}}{l_x}$	Probability of dying within t years

$${}_{n|m}q_x = {}_n p_x \; {}_m q_{x+n}$$

$$D_x = v^x l_x$$

$$C_x = v^{x+1} d_x \qquad\qquad \bar{C}_x = \int_0^1 D_{x+t}\,\mu_{x+t}\,dt$$

$$= v^{-1} D_x - D_{x+1}$$

$$N_x = \sum_{k=0}^{\infty} D_{x+k} \qquad\qquad \bar{N}_x = \int_0^{\infty} D_{x+t}\,dt$$

$$M_x = \sum_{k=0}^{\infty} C_{x+k} \qquad\qquad \bar{M}_x = \int_0^{\infty} D_{x+t}\,\mu_{x+t}\,dt$$

$$= D_x - (1-v^{-1}) N_x$$

$$\ddot{a}_{x:\overline{n}|} = \frac{N_x - N_{x+n}}{D_x} \qquad\qquad \bar{a}_{x:\overline{n}|} = \frac{\bar{N}_x - \bar{N}_{x+n}}{D_x} = \int_0^n v^t \, {}_t p_x \, dt$$

$$A^{1}_{x:\overline{n}|} = \frac{M_x - M_{x+n}}{D_x} \qquad\qquad \bar{A}_{x:\overline{n}|} = \frac{\bar{M}_x - \bar{M}_{x+n}}{D_x} = \int_0^n v^t \, {}_t p_x \, \mu_{x+t}\,dt$$

$$A_{x:\overline{n}|}^{\;\;1} = \frac{D_{x+n}}{D_x} \qquad\qquad \bar{A}_{x:\overline{n}|} = \frac{D_{x+n}}{D_x}$$

$$A_{x:\overline{n}|} = 1 - d\,\ddot{a}_{x:\overline{n}|} \qquad\qquad \bar{A}_{x:\overline{n}|} = 1 - \delta\,\bar{a}_{x:\overline{n}|}$$

Also remember the following relationships from 102

$$i^{(p)} = p\left((1+i)^{1/p} - 1\right) \qquad\qquad \ddot{a}_{x:\overline{n}|} = 1 + a_{x:\overline{n-1}|}$$

$$1+i = v^{-1} = e^{\delta} = \frac{1}{1-d} \qquad\qquad \ddot{a}_{x:\overline{n}|} = v\,p_x\,a_{x+1:\overline{n}|}$$

Most other relationships are built up around an individuals known future lifetime

ω Maximum age anyone may reach (usually 100, 12

$T_x \in [0, \omega]$ An individuals known future lifetime

$K_x = \lfloor T_x \rfloor$ An individuals known lifetime in complete years.

$\mathring{e}_x = E[T_x] = \int_0^{\omega - x} {}_tp_x \, dt$ Complete expectation of life

$e_x = E[K_x] = \sum_0^{\omega - x} {}_tp_x$ Curtite expectation of life

Taking appropriate expectations, we can find values for annuities and assurances

$$\ddot{a}_x = E(\ddot{a}_{\overline{K_x+1|}}) = E\left(\frac{1 - v^{K_x+1}}{d}\right) = \frac{1 - E(v^{K_x+1})}{d} = \frac{1 - A_x}{d}$$

And where we can take expectations, we can take variances. The prefix 2 tells us to calculate using an interest rate of $(1+i)^2 - 1$

$$Var(v^{K_x+1}) = \sum_{k=0}^{\infty} (v^{k+1})^2 {}_{k|}q_x - (A_x)^2 = {}^2A_x - A_x^2$$

$$Var(\ddot{a}_{\overline{K_x+1|}}) = Var\left(\frac{1 - v^{K_x+1}}{d}\right) = \frac{1}{d^2} Var(v^{K_x+1}) = \frac{1}{d^2}({}^2A_x - A_x^2)$$

We can also take the variance of T_x and K_x, but these are less nice.

$$Var[T_x] = \int_0^{\omega - x} t^2 \, {}_tp_x \, \mu_{x+t} - \mathring{e}_x^2$$

$$Var[K_x] = \sum_{k=0}^{\omega - x} k^2 \, {}_kp_x \, q_{x+k} - e_x^2$$

Finally, a lot of relationships arise from the distribution function

$$F_x(t) = P(T_x \leq t) \qquad \text{Distribution function}$$
$$S_x(t) = 1 - F_x(t) \qquad \text{Survival function}$$

$$_tp_x = S_x(t)$$
$$_tq_x = F_x(t)$$
$$\mu_x = \lim_{h \to 0^+} \frac{1}{h} P[T \leq x+h \mid T > x] = \lim_{h \to 0^+} \frac{1}{h} \, {}_hq_x$$

Calculating the PDF gives a formula for $_tq_x$

$$f_x(t) = \frac{d}{dt} F_x(t) = \lim_{h \to 0^+} \frac{1}{h} \left(P(T_x \leq t+h) - P(T_x \leq t) \right)$$
$$= \lim_{h \to 0^+} \frac{1}{S(x)\,h} \left(P(T \leq x+t+h) - P(T \leq x+t) \right)$$
$$= \frac{S(x+t)}{S(x)} \lim_{h \to 0^+} \frac{1}{h} P(T \leq x+t+h \mid T > x+t)$$
$$= {}_tp_x \; \mu_{x+t}$$

$$_tq_x = \int_0^t f_x(s)\,ds = \int_0^t {}_sp_x \, \mu_{x+s} \, ds$$

This can be used to give a formula for $_tp_x$

$$\frac{d}{ds} \, {}_sp_x = -\frac{d}{ds} \, {}_sq_x = -\,{}_sp_x \, \mu_{x+s} \quad \Rightarrow \quad \frac{d}{ds} \log {}_sp_x = -\mu_{x+s}$$

$$_tp_s = e^{-\int_0^t \mu_{x+s}\,ds} \qquad \left({}_0p_x = 1 = e^0 \right)$$

It can also be used to give a nice formula for the central rate of mortality.

$$m_x = \frac{q_x}{\int_0^1 {}_tp_x \, dt} = \frac{\int_0^1 {}_tp_x \cdot \mu_{x+t} \, dt}{\int_0^1 {}_tp_x \, dt}$$

Premiums and Reserves

Insurance companies use a number of terms

Net Premium	Amount of premium required to meet expected cost of benefit
Office Premium	Amount of premium required to meet all costs
Policy Value	The amount a company should set aside to meet future debts
Reserves	The amount a company does set aside.

There are two ways to find policy value - prospective and retrospective. Both wi be equal if they each use the basis used to calculate the premiums.

Prospective policy value = Expected present value of future outgoings
- Expected present value of future income

Retrospective policy value = Accumulated value of premiums received
- Accumulated value of benefits and expenses.

With retrospective values we must allow for interest and survivorship $\left(\frac{D}{D}\right.$ and $\frac{M-M}{D}$

Because the basis has usually changed since the premiums were calculated, we often use the net premium policy value. This is the prospective policy value assuming premiums were calculated on the current basis.

If a person dies, the company looses the death strain at risk $S-({}_{t+1}V+R$
The expected death strain is $q_{x+t}(S-({}_{t+1}V+R))$ and the actual death strain is the observed loss. Mortality profit is defined as EDS - ADS.

A lot of these concepts have a symbolic representation.

$$P_x = \frac{A_x}{\ddot{a}_x} \qquad \text{Net premium for whole life assurance.}$$

$$_tP_{x:\overline{n}|} = \frac{A_{x:\overline{n}|}}{\ddot{a}_{x:\overline{n}|}} \qquad \text{t year premium for n year life assurance.}$$

$$_tV_x = P_x \, \ddot{s}_{x:\overline{t}|} - \frac{(1+i)^t}{_tp_x} A^1_{x:\overline{t}|} \qquad \text{Retrospective policy value}$$

$$_tV_{x:\overline{n}|} = A_{x+t:\overline{n-t}|} - P_{x:\overline{n}|} \, \ddot{a}_{x+t:\overline{n-t}|} \qquad \text{Prospective policy value}$$

$$= 1 - \frac{\ddot{a}_{x+t:\overline{n-t}|}}{\ddot{a}_{x:\overline{n}|}}$$

Using the net premium policy, we find the following relationship.

$$({}_tV_x + P_x)(1+i) = q_{x+t} + p_{x+t} \; {}_{t+1}V_x$$

$$({}_tV_x + P_x)(1+i) = {}_{t+1}V_x + q_{x+t}(1 - {}_{t+1}V)$$

$${}_{t+1}V_x - {}_tV_x = \frac{1}{p_{x+t}} \left(i \, {}_tV_x + (1+i) P_x - q_{x+t}(1 - {}_tV_x) \right)$$

This last equation can be turned into thiels equation for this kind of poli
It's a differential equation you may be asked to form or solve

$$d/dt \; {}_t\bar{V}_x = \delta_t \bar{V}_x + \bar{P}_x - (1 - {}_t\bar{V}_x)\mu_{x+t}$$

Practical Problems

When gathering data to create life tables / calculate policy values, etc. there are a number of practical problems.

Censoring is where we don't know the exact values of each observation. For examp

- Left censoring - we don't know when the observed entered the state of intere
  (Individual falling ill)
- Interval censoring - we only know what interval an event took place in
  (Individual only visits doctor for test once a month)
- Right censoring - observations in progress are cut short
  (Individual leaves, the investigation ends)

- Type I censoring - censoring times known in advance
  (Study continues until 60^{th} birthday)
- Type II censoring - study continues until certain conditions reached
  (Certain number of deaths)

- Non informative / informative censoring - this occurs if T_i and C_i are independant or not respectively. The former is preferred due to easy analysis.

Most methods incorporate a way of dealing with censoring.

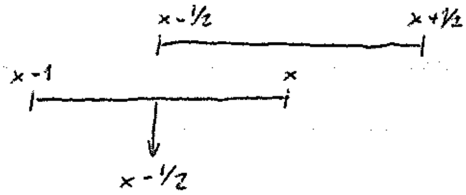
The second problem is that of hetrogeneity. Our assumption that all lives are identical is certainly not the case. Subdividing the data into classes helps this problem at the cost of reducing the population and thus making statistics less credibl

Possible ways to split include age, sex, smoking, type of policy, duration and the level of underwriting.

Another problem is that of dates. How old is someone who was 'x on their nearest birthday at the next policy renewal'?

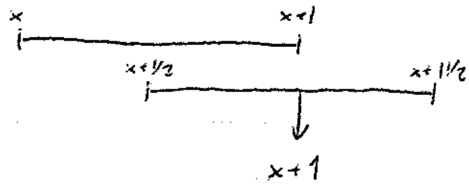
There's no real method here - just work it out logically

x nearest birthday
at the next policy renewal
so they're age $x - \frac{1}{2}$



Similarly with 'age last birthday at entry plus curtate duration'.

x last birthday
at previous anniversary
so they're age $x + 1$



The final problem is how best to integrate between values x and $x+s$ in a life table. There are two main techniques

- Uniform distribution of deaths

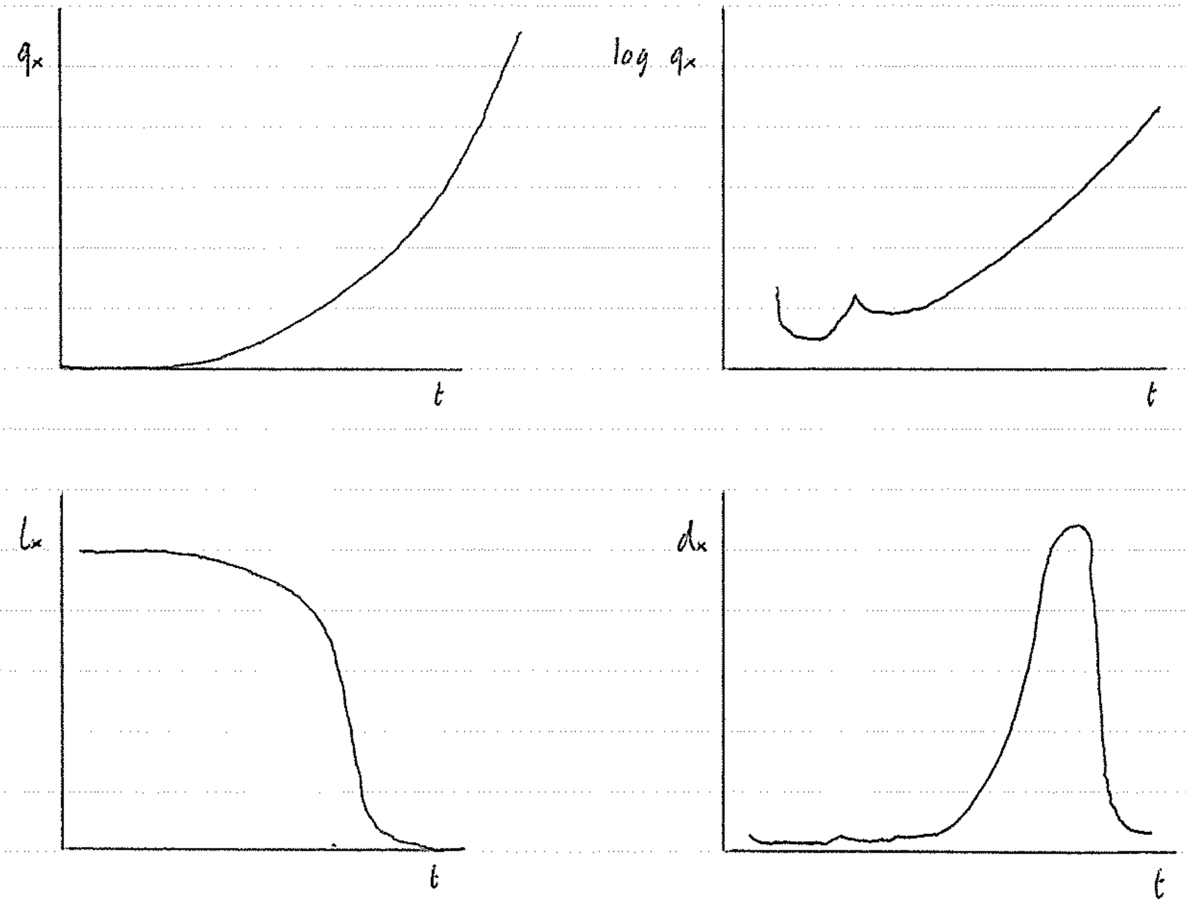
  Assuming ${}_t p_x \, \mu_{x+t}$ $\left(= \frac{q_x}{s}\right)$ is constant we have ${}_{t-s} q_{x-s} = \frac{(t-s)q_x}{1-(s q_x)}$

- Constant force of mortality

  Assuming μ is constant we have ${}_{t-s} p_{x+s} = e^{-(t-s)\mu}$

Pattern of human mortality

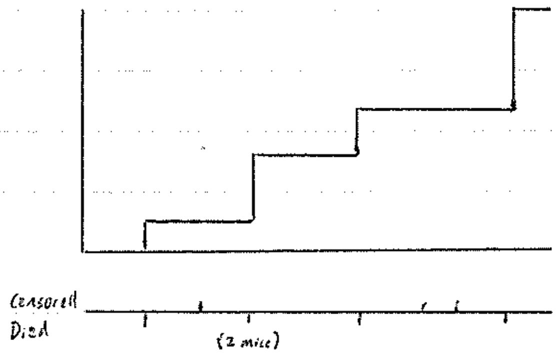
Age	Characteristics	Reasons
0-1	Heavy	Babies born with severe problems
2-12	Light	Protected lifestyle
13-16	Slight increase	Active livestyle
17-20	Temp increase	Accident hump
21-40	Light	Settling down
41-70	Steady increase	Onset of diseases
70+	Heavy, increasing	Body systems lose robustness



Estimating lifetime Distributions

Suppose we have N lives where $d_1 + d_2 + \dots + d_j = m$ deaths occur at times $t_1 < t_2 < \dots < t_j$ and $c_1 + c_2 + \dots + c_j = N - m$ lives were censored between t_i and t_{i+1}. Define n_j as the number of lives alive just before time t_j.

The Kaplan-Meier estimate is the maximum likelyhood estimate for the distribution function. Namely $\hat{F}(t) = 1 - \prod_{t_j \leq t} \left(1 - \frac{d_j}{n_j}\right)$.



d_j/n_j	$1 - d_j/n_j$	$\prod(1 - d_j/n_j)$
$1/8$	$7/8$	$7/8 = 7/8$
$2/6$	$4/6$	$28/48 = 7/12$
$1/4$	$3/4$	$84/192 = 7/16$
$1/1$	$0/1$	$0/192 = 0$

Greenwood's formula estimates the variance $\mathrm{var}(\hat{F}(t)) = (1 - \hat{F}(t))^2 \sum_{t_j \leq t} \frac{d_j}{n_j(n_j - d_j)}$, but may understate the variance of the tails a little.

The Nelson-Aalen estimate is a little more sophisticated $\hat{F}(t) = 1 - e^{-\sum_{t_j \leq t} \frac{d_j}{n_j}}$ and approximates the Kaplan-Meier estimate.

Greenwood's formula estimates the variance $\mathrm{var}\left(-\sum_{t_j \leq t} \frac{d_j}{n_j}\right) \simeq \sum_{t_j \leq t} \frac{n_j - d_j}{n_j^3}$ for the variance of the integrated hazard.

The Cox Model

The Cox model has a hazard function $\lambda(t; z_i) = \lambda_0(t) \exp(\beta z_i^T)$.
The baseline hazard $\lambda_0(t)$ describes the hazard for the null covariate vector $z_i = \underline{0}$.
This is in/decreased proportionally by $\exp(\beta z_i^T)$.

The usefulness of this model arises from the fact that if we are interested only in the covariate effect, we can completely ignore the baseline hazard.

For example, to estimate β, we maximise the partial likelyhood $L(\beta) = \prod_{j=0}^{k} \frac{\exp(\beta z_j^T)}{(\sum_{i \in R(t_j)} \exp(\beta z_i$
The baseline hazard cancels out and we are only concerned with the order.
For simultaneous deaths use the approximation $L(\beta) = \prod_{j=1}^{k} \frac{\exp(\beta \Sigma z_j^T)}{(\sum_{i \in R(t_j)} \exp(\beta z_i^T))^{d_j}}$
To solve we find $u(\beta) = \left(\frac{\partial \log L(\beta)}{\partial \beta_1}, \dots, \frac{\partial \log L(\beta)}{\partial \beta_p} \right) = \underline{0}$.

Suppose we need to assess the effect of adding covariates to the model.
Then if L_p, L_{p+q} are maximised log likelyhoods of models with $p, p+q$ covariates, under the hypothesis that covariates have no effect $-2(L_p - L_{p+q}) \sim \chi^2_q$

The Markov Model

The markov model has three assumptions

- The probability of a future state depends only on the current state and transition probabili
- For two states g,h ${}_{dt}p^{gh}_{x+t} = \mu^{gh}_{x+t}\,dt + o(dt)$. Multiple transitions are $o(dt)$
- μ_{x+t} is constant μ for the period considered.

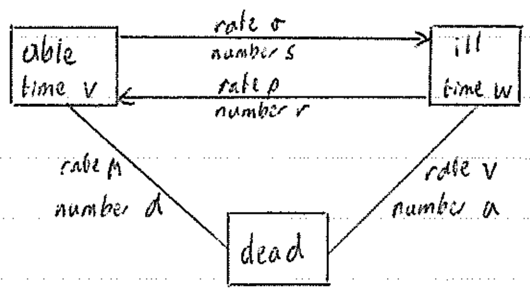
The main equations are the Kolmogorov forward equations

$$\frac{d}{dt}\,{}_tp_x^{gh} = \sum_{j\neq h}\left({}_tp_x^{gj}\,\mu_{x+t}^{jh} - {}_tp_x^{gh}\,\mu_{x+t}^{hj}\right)$$

and that

$$\frac{d}{dt}\,{}_tp_x^{\overline{gg}} = -\,{}_tp_x^{\overline{gg}}\sum_{j\neq g}\mu_{x+t}^{gj} \quad\Rightarrow\quad {}_tp_x^{\overline{gg}} = e^{-\int_0^t \sum_{j\neq g}\mu_{x+s}^{gj}\,ds}$$

The other thing we may have to do is estimate transition parameters using MLEs. For example, if we have the following model...



The MLE is $L(\mu,\nu,\sigma,\rho) = e^{-(\mu+\sigma)v}\,e^{-(\nu+\rho)w}\,\mu^{d}\,\nu^{u}\,\sigma^{s}\,\rho^{r}$ which can be solved to give $\tilde{\mu} = D/V$, $\tilde{\nu} = U/W$, $\tilde{\sigma} = S/V$, $\tilde{\rho} = R/W$.

These estimators are not independant but are asymptotically independant. They also form a multivariate normal distribution with individual components $\tilde{\mu} \sim N\left(\mu, \frac{\mu}{E(V)}\right)$.

Miscellanious Models

The following two models rely simply on the exponential nature of death

Gompertz Law $\mu_x = Bc^x$
Makeham's Law $\mu_x = A + Bc^x$

The binomial model is another simple model used to estimate the number dying in a year. Observing N lives aged x, we get a sample d of the deaths D. Then $D \sim Bin(N, \hat{q}_x = \frac{d}{N})$.

This has problems with lives only observed between a_i and b_i. We solve this by taking two vectors $\vec{q} = ({}_{b_i - a_i} q_{x+a_i})$ and $\vec{d} = (I_{\{\text{subject dies}\}})_i$ so

$$L(\vec{q};\vec{d}) = \prod {}_{b_i-a_i} q_{x+a_i}^{\ d_i} \left(1 - {}_{b_i-a_i} q_{x+a_i}\right)^{1-d_i}$$

Estimating ${}_{1-t} q_{x+t} = (1-t) q_x$, we find $\vec{q}$ which maximises L. Now as

$$E(D) = \left(\sum (1-a_i) - \sum_{i:D_i=0} (1-b_i)\right) q_x$$

$$\Rightarrow \hat{q}_x = \frac{d}{\sum (1-a_i) - \sum_{i:d_i=0} (1-b_i)} \approx \frac{d}{E_x^c + d/2} \quad \text{(the actuarial estimate)}$$

The poisson distribution is given by $P(D=d) = \frac{e^{-\mu E_x^c} (\mu E_x^c)^d}{d!}$
We get the following estimator of $\tilde{\mu} = \frac{D}{E_x^c}$ which matches the 2 state model

When choosing models, the two state, then the poisson, then binomial are best. This is in terms of modelling the process, finding parameters and extending models. The one area the two state model falls down is that its MLE is only asymptotically unbiased, and the variance can only be found asymptotically, while its immediately availible in poisson and binomial. However, for small μ, differences are minor in all these respects.

Methods of Graduation

We wish to graduate life tables because

- We feel they should be smooth
- It uses data at adjacent ages to improve the estimate
- Smoothness is important for financial products.

However, when smoothing we must balance it against following the data, and must ensure that estimates do not take us in unsuitable directions (ie cause a los

There are three main methods

Fitting to a formula - choose the formula, get parameter values, calculate the rates and test for suitability.

Fitting to a standard table - choose the table, choose a relationship between our values and it, estimate parameter values, calculate the rates and test for suitability.

Fitting graphically - Plot estimate, mark confidence intervals, sketch graph read off rates and test for suitability.

Fitting to formulae is good for comparisons as you can compare parameters, good for creating standard tables with lots of data.

Fitting to a standard table is good for fitting scarce data where we know an appropriate table. It allows us to extrapolate well, where there may be little dat

Fitting graphically allows an experienced person to allow for known features wi scanty data. On the other hand any bias will be equally reflected.

Graduation Tests

Firstly, there are a number of things that we should check immediately

- Mortality of males greater than females
- Mortality of people with life insurance lower than the population.

Secondly, we want to check for smoothness. We do this by taking the third difference and ensuring this is low and progresses regularly.

Thirdly, there are a number of tests.

- Signs test (if graduated rates are too high or low)

  Count the number of grad rates above crude rates (n_1).
  This follows Bin (m, $\frac{1}{2}$). Perform a two tailed test.

- Grouping of Signs (detects clumping of same sign deviations)

  Find the smallest k such that $\sum_{n=1}^{k} \frac{\binom{n_1-1}{t-1}\binom{n_2+1}{t}}{\binom{m}{n_1}} \geq 0.05$.
  The test fails if the number of groups $G \leq k$.

  For large m, we have $G \sim N\left(\frac{n_1(n_2+1)}{m}, \frac{(n_1 n_2)^2}{m^3}\right)$

- Serial Correlations test (detects clumping of same sign deviations)

  The null hypothesis is that $z_1, \ldots, z_{m-1}$ and $z_2, \ldots, z_m$ should be uncorrelated.

$$\sqrt{m}\frac{\sum (z_i - \bar{z}_0)(z_{i+j} - \bar{z}_j)}{\sqrt{\sum (z_i - \bar{z}_0)^2 \sum (z_{i+j} - \bar{z}_j)^2}} \simeq \sqrt{m}\frac{\sum (z_i - \bar{z})(z_{i+j} - \bar{z})}{\frac{m-j}{m}\sum (z_i - \bar{z})^2} \sim N(0,1)$$

  If this is too high, reject the null hypothesis.

Let $z_x = \frac{d_x - E_x^c \mu^s_{x+1/2}}{\sqrt{E_x^c \mu^s_{x+1/2}}}$ (poisson)

$= \frac{d_x - E_x^c q_x^s}{\sqrt{E_x^c q_x^s (1-q_x^s)}}$ (binomial)

- Chi^2 Test (test for fit)

Pool groups so that each one has > 5 members. and calculate $\sum_{\text{each grp}} z_x^2$
Calculate the degrees of freedom. - if comparing against a standard table, the answer is m, if against a graduation (m - no. parameters).

- Standardised Deviations Test (test for over/under graduation)

The z_x s should form $N(0,1)$. Either use a χ^2 test to see if they do, or visually inspect them. Splits are 0.02, 0.14, 0.34, 0.34, 0.14, 0.02.

- Cumulative Deviation (Goodness of fit for small areas)

Pick a range. Then $\frac{\sum (d_x - E_x^c \mu_{x+1/2})}{\sqrt{\sum E_x^c \mu_{x+1/2}}} \sim N(0,1)$