ETC3430/ETC5343
Financial Mathematics under Uncertainty

Dan Zhu

Monash Business School

March 23, 2022

You need some (basic) knowledge about:

- maximum likelihood estimation methods
- law of large numbers
- asymptotic distribution and central limit theorem
- asymptotic properties of maximum likelihood estimator

Models and Model parameters

We use a model to describe the process that results in the data that are observed. For example, we may use a linear model to predict the revenue that will be generated for a company depending on how much they may spend on advertising (this would be an example of linear regression)

$$y_i = x_i\beta + \epsilon_i \ \epsilon \sim N(0, \sigma^2)$$

Model Specification: Linear, Gaussian
Model Parameters: β, σ^2
Each model contains its own set of parameters that ultimately defines what the model looks like.

Maximum Likelihood Estimate

Maximum likelihood estimation is a method that determines values for the parameters of a model. The parameter values are found such that they maximise the likelihood that the process described by the model produced the data that were actually observed.

Input:

- Data: independent random variables $X_1, \ldots, X_n$ with observed samples: $x_1, \ldots, x_n$
- Model: the same probability density/mass function $f(x; \theta)$, i.e., the model we think best describes the process of generating the data that usually comes from having some domain expertise but we wont discuss this here.

Objective: Estimate $\theta = (\theta_1, \ldots, \theta_p)$, some **unknown** parameters.

MLE continued

MLE utilitse the likelihood function (fitness of the parameters to the sample)

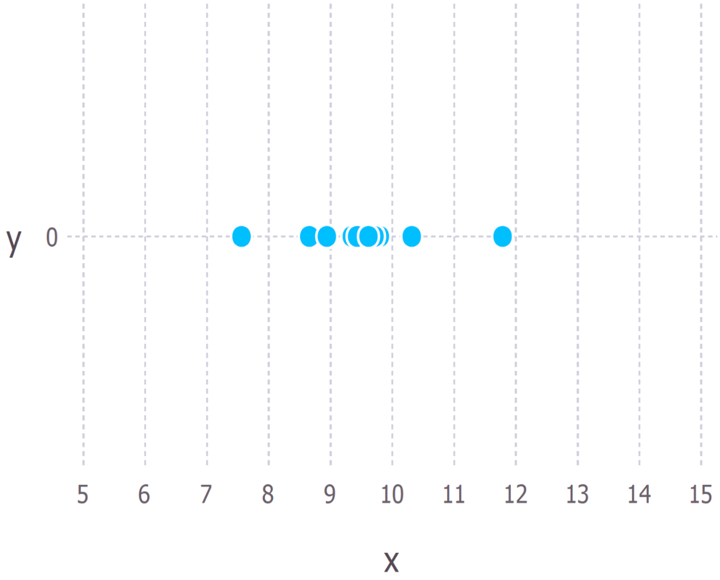
$$\mathcal{L}(\theta; x_1, \ldots, x_n) = f(x_1; \theta) \cdot f(x_2; \theta) \cdots f(x_n; \theta)$$

or simply $\mathcal{L}(\theta)$. maximum likelihood estimate

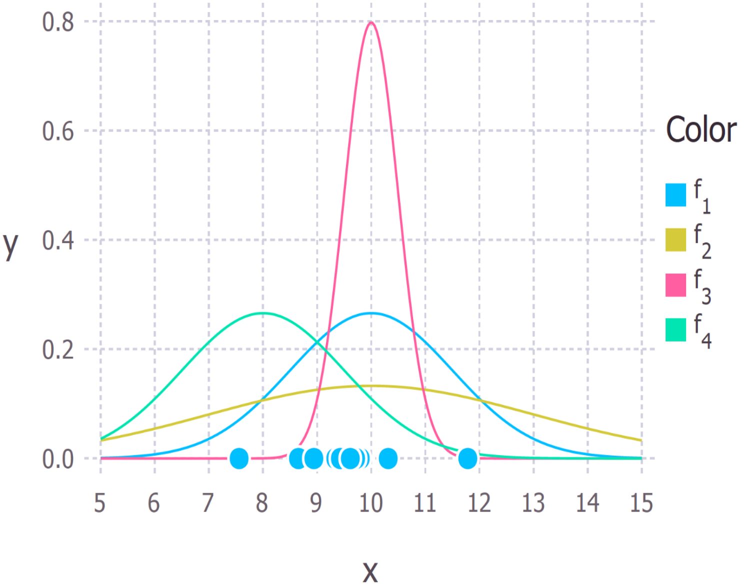
$$\widehat{\theta} = \arg\max_{\theta} \mathcal{L}(\theta) = \arg\max_{\theta} \log \mathcal{L}(\theta)$$

$\widehat{\theta} = \widehat{\theta}(x_1, \ldots, x_n)$ depends on the observations $x_1, \ldots, x_n$

Observe Some data



Which Gaussian



Asymptotic Distribution: ML Estimator

The maximum likelihood **estimator**

$$\widetilde{\theta} = \widehat{\theta}(X_1, \ldots, X_n)$$

is a random variable itself because it is a function of the $X_1, ..., X_n$.

- exact distribution is hard to compute (simulation?)
- **asymptotic** distribution: limiting distribution as $n \to \infty$

  $$N(\theta_0, \mathcal{I}^{-1})$$

  where

  $$\mathcal{I}_{ij} = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log \mathcal{L}(\theta; X_1, \ldots, X_n)\right)$$

  denote by $\widetilde{\theta} \overset{a}{\sim} \mathcal{N}\left(\theta, \mathcal{I}^{-1}\right)$
- Large n: exact distribution $\approx$ asymptotic distribution

Law of large numbers

Theorem

If $X_1, \ldots, X_n$ is a sequence of independent and identically distributed random variables with finite mean μ, then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu, \quad \text{as } n \to \infty$$

with probability 1.

- weak law of large numbers: $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$
- heuristic: $\frac{1}{n}\sum_{i=1}^{n} X_i \approx \mu$

UnivariateCentral limit theorem

Theorem

Univariate central limit theorem If $X_1, \ldots, X_n$ is a sequence of independent and identically distributed random variables with finite mean μ and variance σ^2,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \overset{a}{\sim} \mathcal{N}\left(\mu, \sigma^2\right), \text{ as } n \to \infty.$$

Multivariate central limit theorem

Theorem

If $X_1, \ldots, X_n$ is a sequence of independent and identically distributed random vectors with finite mean μ and covariance matrix Σ, as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \overset{a}{\sim} \mathcal{N}(\mu, \Sigma), \text{ as } n \to \infty.$$

An interesting thing about the CLT is that it does not matter what the distribution of the X_i's is. It can be discrete, continuous, or mixed random variables.

Example

$X_i \sim Bernoulli(p)$, then

$$\mathbb{E}[X_i] = p, \; and \; var(X_i) = p(1-p).$$

Assumptions:

- $X_1, X_2 \ldots$ are iid Bernoulli(p).
- $Z_n = \dfrac{X_1 + X_2 + \ldots + X_n - np}{\sqrt{np(1-p)}}$.

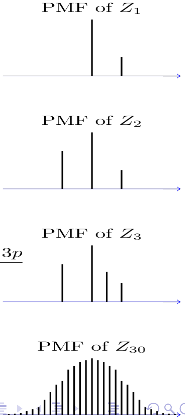
We choose $p = \frac{1}{3}$.

$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$



Example

$X_i \sim Uniform(0,1)$, then

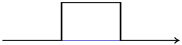
$$\mathbb{E}[X_i] = 0.5, \ \textit{and } var(X_i) = \frac{1}{12}.$$

Assumptions:

- $X_1, X_2 \ldots$ are iid Uniform(0,1).
- $Z_n = \dfrac{X_1 + X_2 + \ldots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$.

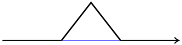
$$Z_1 = \frac{X_1 - \frac{1}{2}}{\sqrt{\frac{1}{12}}}$$

PDF of Z_1



$$Z_2 = \frac{X_1 + X_2 - 1}{\sqrt{\frac{2}{12}}}$$

PDF of Z_2



$$Z_3 = \frac{X_1 + X_2 + X_3 - \frac{2}{3}}{\sqrt{\frac{3}{12}}}$$

PDF of Z_3

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - \frac{30}{2}}{\sqrt{\frac{30}{12}}}$$

PDF of Z_{30}

Some asymptotic properties

Slutsky's theorem

Let $\widetilde{\theta}_1 \overset{a}{\sim} \mathcal{N}(\theta_1, \sigma_1^2)$ and $\widetilde{\theta}_2 \approx c$.

- $\widetilde{\theta} = \widetilde{\theta}_1 - \widetilde{\theta}_2 \overset{a}{\sim} \mathcal{N}(\theta_1 - c, \sigma_1^2)$
- $\widetilde{\theta} = \widetilde{\theta}_1 \cdot \widetilde{\theta}_2 \overset{a}{\sim} \mathcal{N}\left(c\theta_1, c^2\sigma_1^2\right)$
- $\widetilde{\theta} = \frac{\widetilde{\theta}_1}{\widetilde{\theta}_2} \overset{a}{\sim} \mathcal{N}\left(\frac{\theta_1}{c}, \frac{\sigma_1^2}{c^2}\right)$ if $c \neq 0$

Asymptotic Independence

$$\begin{pmatrix} \widetilde{\theta}_1 \\ \widetilde{\theta}_2 \end{pmatrix} \overset{a}{\sim} \mathcal{N}\left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right)$$

then $\widetilde{\theta}_1$ and $\widetilde{\theta}_2$ are asymptotically independent.

Confidence Interval

An interval $\left(\widetilde{\theta}_L(x_1, \ldots, x_n), \widetilde{\theta}_U(x_1, \ldots, x_n)\right)$ is

- a 95% confidence interval for θ if

$$\mathbb{P}\left(\widetilde{\theta}_L(X_1, \ldots, X_n) < \theta < \widetilde{\theta}_U(X_1, \ldots, X_n)\right) = 95\%$$

- a 95% **asymptotic** confidence interval for θ if

$$\mathbb{P}\left(\widetilde{\theta}_L(X_1, \ldots, X_n) < \theta < \widetilde{\theta}_U(X_1, \ldots, X_n)\right) \to 95\%, \; n \to \infty.$$

  For large n,

$$\mathbb{P}\left(\widetilde{\theta}_L(X_1, \ldots, X_n) < \theta < \widetilde{\theta}_U(X_1, \ldots, X_n)\right) \approx 95\%$$

Confidence Interval

Suppose $\widetilde{\theta} = \widetilde{\theta}(X_1, \ldots, X_n) \overset{a}{\sim} \mathcal{N}\left(\theta, \sigma_n^2\right)$:

- estimate $\widehat{\theta} = \widetilde{\theta}(x_1, \ldots, x_n)$
- an 95% asymptotic confidence interval is

$$\left(\widehat{\theta} - 1.96\sigma_n, \widehat{\theta} + 1.96\sigma_n\right), \text{ or } \quad \widehat{\theta} \pm 1.96\sigma_n$$

  where 1.96 is rounded from 1.95996 . . ..
- if σ_n is unknown but estimable by $\widehat{\sigma}_n$, use the interval

$$\left(\widehat{\theta} - 1.96\widehat{\sigma}_n, \widehat{\theta} + 1.96\widehat{\sigma}_n\right), \text{ or } \quad \widehat{\theta} \pm 1.96\widehat{\sigma}_n.$$

A Bernoulli Example

In fact, the Bernoulli distribution $B(p)$ is not a single distribution but rather a one-parameter family of distributions. Each value of p defines a different distribution in the family, with the pdf

$$f(x) = p^x(1-p)^1 - x \; x \in \{0, 1\}.$$

There are many methods for estimating unknown parameters from data. We will first consider the maximum likelihood estimate (MLE), which answers the question: For which parameter value does the observed data have the most significant probability?

Example Continue

A coin is flipped 100 times. Given that there were 55 heads find the maximum likelihood estimate for the probability p of heads on a single toss.

- Experiment: Flip the coin 100 times and count the number of heads.
- Data: The data is the result of the experiment. In this case, it is 55 heads
- Parameter of interest: We are interested in the value of the unknown parameter p.
- Likelihood, or likelihood function

$$L_1 00(p) = \binom{100}{55} p^{55}(1-p)^{100-55}$$

Optimal $\hat{p} = \frac{55}{100} = 0.55$.

Two States Markov Chain

For lives from age x to $x+1$ obeying a Markov model with two states: alive and dead

$$\text{alive} \xrightarrow{\mu} \text{dead}$$

Assumption

The mortality rate, i.e. transition rate, $\mu_{x+t} = \mu$ for all $t \in [0, 1)$

For $t \in [0, 1)$,

$$\begin{aligned} {}_tp_x =& P\left(\text{Alive at age } t+x | \text{Alive at age } x\right) \\ =& \exp\left(-\int_0^t \mu_{x+s} ds\right) = \exp\left(-\mu t\right) \end{aligned}$$

Estimating the Mortality rate

- the value of μ is often unknown in real life
- μ may be estimated from observations

Data of individuals

- observations of n lives indexed by $i \in \{1, \ldots, n\}$
- starting from $x + a_i$ and ends no later than $x + b_i$
- censored (more in next week): $0 \leq a_i < b_i \leq 1$
- i-th life survives to $x + T_i$ in observation
- waiting time $V_i \equiv T_i - a_i \in (0, b_i - a_i]$
- $D_i = 1$ if i-th life ends during observation otherwise $D_i = 0$
- our dataset is $\{(v_i, d_i) : i = 1, \ldots, n\}$

Likelihood

Assumption

The lives are identical and statistically independent, i.e., $\{(V_i, D_i) : i = 1, \ldots, n\}$ are i.i.d. distributed.

- D_i fully depends on V_i: $D_i = 0 \Leftrightarrow V_i = b_i - a_i$
- V_i has a probability mass

$$\mathbb{P}(V_i = b_i - a_i) = {}_{b_i - a_i}\, p_x = \exp\left(-\mu(b_i - a_i)\right)$$

- ... and density function on $(0, b_i - a_i)$

$$\begin{aligned} f(t; \mu) =& \frac{\partial}{\partial t} {}_t q_x \\ =& \frac{\partial}{\partial t}\left(1 - \exp\left(-\mu t\right)\right) = \mu \exp\left(-\mu t\right), \quad t \in (0, b_i - a_i) \end{aligned}$$

- Likelihood for $i \in \mathcal{C} = \{i : d_i = 0\}$

$$\mathcal{L}_i(\mu) = \exp(-\mu(b_i - a_i)) = \exp(-\mu v_i) = \mu^{d_i} \exp(-\mu v_i)$$

- Likelihood for $i \in \mathcal{D} = \{i : d_i = 1\}$

$$\mathcal{L}_i(\mu) = f(v_i; \mu) = \mu \exp(-\mu v_i) = \mu^{d_i} \exp(-\mu v_i)$$

Total likelihood

$$\mathcal{L}(\mu) = \Pi_{i=1}^{n} \mathcal{L}_i(\mu) = \mu^{d} \exp(-\mu v)$$

where

$$d = \sum_{i=1}^{n} d_i, \quad v = \sum_{i=1}^{n} v_i$$

Maximum Likelihood Estimation

- $\log \mathcal{L}(\mu) = d \log \mu - \mu v$
- $(\log \mathcal{L}(\mu))' = \frac{d}{\mu} - v = 0$ gives $\widehat{\mu}_x = \frac{d}{v}$
- $(\log \mathcal{L}(\mu))'' = -\frac{d}{\mu^2} < 0$

The maximum likelihood estimate is

$$\widehat{\mu} = \frac{d}{v} = \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} v_i}$$

and the maximum likelihood estimator

$$\widetilde{\mu} = \frac{D}{V} = \frac{\sum_{i=1}^{n} D_i}{\sum_{i=1}^{n} V_i}$$

- average number of deaths from the population for each time unit

Asymptotic Properties of $\widetilde{\mu}$

$$\widetilde{\mu} - \mu = \frac{\frac{1}{n}\sum_{i=1}^{n}(D_i - \mu V_i)}{\frac{1}{n}\sum_{i=1}^{n} V_i}$$

- Homeworks:

$$\mathbb{E}(D_i - \mu V_i) = 0, \ \mathrm{Var}(D_i - \mu V_i) = \mathbb{E}(D_i - \mu V_i)^2 = \mu E(V_i)$$

- Central limit theorem:

$$\frac{1}{n}\sum_{i=1}^{n}(D_i - \mu V_i) \overset{a}{\sim} \mathcal{N}\left(0, \frac{1}{n}\mu E(V_i)\right)$$

- Law of large numbers: $\frac{1}{n}\sum_{i=1}^{n} V_i \approx E(V_i)$
- Slutsky theorem:

$$\widetilde{\mu} - \mu \overset{a}{\sim} \mathcal{N}\left(0, \frac{1}{n}\frac{\mu E(V_i)}{(E(V_i))^2}\right)$$

From the last step:

$$\widetilde{\mu} - \mu \overset{a}{\sim} \mathcal{N}\left(0, \frac{1}{n}\frac{\mu}{E\left(V_i\right)}\right)$$

or equivalently

$$\widetilde{\mu} \overset{a}{\sim} \mathcal{N}\left(\mu, \frac{\mu}{E\left(V\right)}\right)$$

since $E(V) = \sum_{i=1}^{n} E(V_i) = nE(V_i)$

95% Confidence Interval

From the last step:

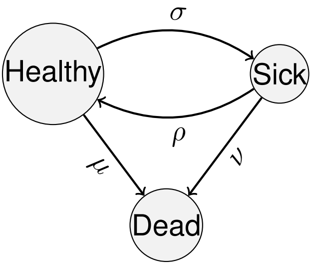
$$\widehat{\mu} \pm 1.96\sqrt{\frac{\mu}{E(V)}}$$

Estimating μ by $\widehat{\mu}$ and $E(V)$ by v yields the asymptotic 95% confidence interval

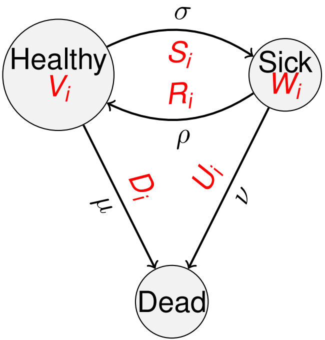
$$\widehat{\mu} \pm 1.96\sqrt{\frac{\widehat{\mu}}{v}}$$

HSD Model

For lives from obeying a Markov model with three states: healthy (H), sick (S) and dead (D)



Data



the i-th life:
V_i=Total waiting time in 'healthy'
W_i=Total waiting time in 'sick'
S_i=# transitions 'healthy' to 'sick'
R_i=# transitions 'sick' to 'healthy'
D_i=# transitions 'healthy' to 'dead'
U_i=# transitions 'sick' to 'dead'

Observed: v_i, w_i, s_i, r_i, d_i and u_i.
$v_i = \sum_{j=1}^{d_i+s_i} v_{i,j}$ and $w_i = \sum_{j=1}^{r_i+u_i} w_{i,j}$

Individual likelihood

Last week, Page 3

Given a jump has occurred, the time at which took place does affect the probability of the jump being to a particular state.

Two independent components:

1. waiting times
   - at state 'Healthy':

     $$\Pi_{j=1}^{d_i+s_i}(\mu+\sigma)\exp\left(-(\mu+\sigma)v_{i,j}\right) = (\mu+\sigma)^{d_i+s_i}\exp\left(-(\mu+\sigma)v_i\right)$$

   - at state 'Sick'

     $$\Pi_{j=1}^{r_i+u_i}(\rho+\nu)\exp\left(-(\rho+\nu)w_{i,j}\right) = (\rho+\nu)^{r_i+u_i}\exp\left(-(\rho+\nu)w_i\right)$$

2. transitions between states

Individual likelihood: continued

Two independent components:

1. waiting times
   - at state 'Healthy': $(\mu+\sigma)^{d_i+s_i}\exp\left(-(\mu+\sigma)\,v_i\right)$
   - at state 'Sick': $(\rho+\nu)^{r_i+u_i}\exp\left(-(\rho+\nu)\,w_i\right)$
2. transitions between states
   - $H\to S$: $\left(\frac{\sigma}{\mu+\sigma}\right)^{s_i}$
   - $H\to D$: $\left(\frac{\mu}{\mu+\sigma}\right)^{d_i}$
   - $S\to H$: $\left(\frac{\rho}{\rho+\nu}\right)^{r_i}$
   - $S\to D$: $\left(\frac{\nu}{\rho+\nu}\right)^{u_i}$

Multiplying all components,

$$\mathcal{L}_i\left(\mu,\nu,\sigma,\rho\right)=\exp\left(-(\mu+\sigma)\,v_i\right)\exp\left(-(\rho+\nu)\,w_i\right)\mu^{d_i}\nu^{u_i}\sigma^{s_i}\rho^{r_i}$$

Total likelihood

Recall from the last step:

$$\mathcal{L}_i(\mu,\nu,\sigma,\rho) = \exp\left(-(\mu+\sigma)v_i\right)\exp\left(-(\rho+\nu)w_i\right)\mu^{d_i}\nu^{u_i}\sigma^{s_i}\rho^{r_i}$$

The individuals are independent:

$$\begin{aligned}\mathcal{L}(\mu,\nu,\sigma,\rho) &= \Pi_{i=1}^{n}\mathcal{L}_i(\mu,\nu,\sigma,\rho)\\ &= \exp\left(-(\mu+\sigma)v\right)\exp\left(-(\rho+\nu)w\right)\mu^{d}\nu^{u}\sigma^{s}\rho^{r}\end{aligned}$$

where

$$v=\sum_{i=1}^{n}v_i,\ w=\sum_{i=1}^{n}w_i,\ s=\sum_{i=1}^{n}s_i,$$
$$r=\sum_{i=1}^{n}r_i,\ d=\sum_{i=1}^{n}d_i,\ u=\sum_{i=1}^{n}u_i.$$

Maximum Likelihood

$$\log \mathcal{L} = -(\mu+\sigma)\,v - (\rho+\nu)\,w + d\log\mu + u\log\nu + s\log\sigma + r\log\rho$$

$$\frac{\partial\mathcal{L}}{\partial\mu} = -v + \frac{d}{\mu},\ \frac{\partial\mathcal{L}}{\partial\nu} = -w + \frac{u}{\nu},\ \frac{\partial\mathcal{L}}{\partial\sigma} = -v + \frac{s}{\sigma},\ \frac{\partial\mathcal{L}}{\partial\rho} = -w + \frac{r}{\rho}$$

Setting all partial derivatives to be zero, yields the maximum likelihood estimates

$$\widehat{\mu} = \frac{d}{v},\ \widehat{\nu} = \frac{u}{w},\ \widehat{\sigma} = \frac{s}{v},\ \widehat{\rho} = \frac{r}{w}$$

They are maximum points because the log-likelihood function is concave: the Hessian matrix

$$\mathbf{H}(\mu,\nu,\sigma,\rho) = \mathrm{diag}\left(-\frac{d}{v^2}, -\frac{u}{v^2}, -\frac{s}{\sigma^2}, -\frac{r}{\rho^2}\right)$$

is negative definite.

Maximum likelihood estimates in general

The maximum likelihood estimate of the transition rate μ_{km} from state k to state m, $k \neq m$, is

$$\widehat{\mu}_{km} = \frac{n_{km}}{t_k} = \frac{\text{\# transitions from state } k \text{ to state } m}{\text{total waiting time in state } k}$$

- This is a general formula for time-homogeneous Markov jump process with any finite number of states.
- Denote by the corresponding estimator to be $\widetilde{\mu}_{km}$.

Asymptotic Normality

K-states time-homogeneous Markov jump process:

- estimators $\{\widetilde{\mu}_{km} : k, m \in \{1, \ldots, K\}, k \neq m\}$ are asymptotically multivariate normal
- ...and asymptotically independent
  - The Hessian matrix of log-likelihood function is diagonal, and so does the Fisher information;
  - HSD model: $D_i - \mu V_i$, $U_i - \mu V_i$, $S_i - \sigma V_i$, and $R_i - \rho W_i$ are uncorrelated

Confidence Intervals

just like in the two-states model:

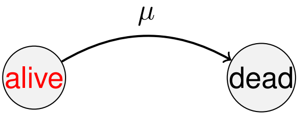
- $$\widetilde{\mu}_{km} - \mu_{km} \overset{a}{\sim} \mathcal{N}\left(0, \frac{\mu_{km}}{\mathbb{E}(T_k)}\right)$$

  where T_k is the total waiting time in state k over all individuals.
- asymptotic 95% confidence intervals for μ_{km}

$$\widehat{\mu}_{km} \pm 1.96\sqrt{\frac{\widehat{\mu}_{km}}{t_k}}$$

Estimating from life tables

Recall the alive and dead model for lives from age x to $x+1$:



- Obtaining total waiting time $v = \sum_{i=1} v_i$ may be infeasible or too costly
- however, the number of deaths $d = \sum_{i=1}^{n} d_i$
- ...and the population of different age groups may be obtained from census data
- can we estimate μ?

Expected number of deaths

- $D_i = 1$ is a Bernoulli variable

$$\mathbb{E}(D_i) = \mathbb{P}(D_i = 1) = 1 - e^{-\mu(b_i - a_i)} \approx \mu(b_i - a_i)$$

- The expected number of deaths

$$\mathbb{E}(D) = \sum_{i=1}^{n} \mathbb{E}(D_i) \approx \mu \sum_{i=1}^{n} (b_i - a_i) = \mu E_x^c$$

- ...where $E_x^c = \sum_{i=1}^{n} (b_i - a_i)$ is called central exposed to risk (more in Week 8)

Poisson Model

Theorem

Poisson limit theorem: when D_i are independent and under some weak regularity conditions,

$$D = \sum_{i=1}^{n} D_i \overset{a}{\sim} Poisson(\lambda)$$

where $\lambda = \mu E_x^c$.

The Poisson approximation:

$$\mathbb{P}(D = d) \approx \frac{e^{-\mu E_x^c} (\mu E_x^c)^d}{d!}$$

the approximation allows the impossible event $D > n$ with a small probability

MLE in Poisson Model

Now use the approximate Poisson likelihood

$$\mathcal{L}(\mu) = \frac{e^{-\mu E_x^c} \left(\mu E_x^c\right)^d}{d!}$$

Taking logarithms of both sides yields that

$$\log \mathcal{L}(\mu) = -\mu E_x^c + d \left(\log(\mu) + \log(E_x^c)\right) - \log(d!)$$

Solve the first order condition

$$(\log \mathcal{L}(\mu))' = -E_x^c + \frac{d}{\mu} = 0$$

and we obtain the maximum likelihood estimator

$$\widehat{\mu} = \frac{d}{E_x^c}$$

- Second order condition: $(\log \mathcal{L}(\mu))'' = -\frac{d}{\mu^2} < 0$.

Maximum Likelihood Estimator in Poisson Model

The maximum likelihood estimator is therefore

$$\widetilde{\mu} = \frac{D}{E_x^c}$$

and the maximum likelihood estimate is

$$\widehat{\mu} = \frac{d}{E_x^c}$$

- Substituting E_x^c for V in the previous MLE $\widetilde{\mu} = \frac{D}{V}$

Asymptotic Property of $\widetilde{\mu}$

- When $E(D) \approx \mu E_x^c$ is large, we can use the normal approximation

$$D \overset{a}{\sim} \mathcal{N}\left(\mu E_x^c, \mu E_x^c\right)$$

and thus

$$\widetilde{\mu} = \frac{D}{E_x^c} \overset{a}{\sim} \mathcal{N}\left(\mu, \frac{\mu}{E_x^c}\right)$$

- 95% asymptotic confidence interval

$$\widehat{\mu} \pm 1.96\sqrt{\frac{\widehat{\mu}}{E_x^c}}$$

- same formula as before except substituting E_x^c for v