

# Child survival for mothers: mortality change and related measures

Ivan Williams and Diego Alburez-Gutierrez

07 August 2019

## Introduction

Kin count estimation has a long pedigree in mathematical demography. Lotka (1931) proposed to model the orphanhood in a theoretical population, comparing the effect across demographic regimes. Equation ?? was originally proposed by Brass (1953) to estimate child mortality. The notion of estimating the expected number of living daughters in a stable population was generalized by Goodman, Keyfitz, and Pullum (1974) for other kin relations (granddaughters, cousins, etc.). The “counting method” approach, being the net reproduction rate  $R_0$  limited to  $a$ , or the (not complete) ratio between generations (multiplying and dividing by the mother generation  $B$ ), was further popularized by Keyfitz and Caswell (2005). Bongaarts (1987) used a similar approach to estimate descendants in his ‘Family Status Model’. More recently, Wrycza and Baudisch (2012) has looked at the formal effect of different kind of changes in mortality by age.

The objective of this work is to give a formal approach in an stable regime context to the timing of descendent mortality (mean age at the event and time spent in that state, absolute and relative) and the effect of mortality change. Finally, a numerical application of these expressions will be done for Latinoamerican countries during the second half of XX Century, relaxing also the stable assumption with an application.

## Method

$CS_a$  is the expected number of surviving children to a woman alive at aged  $a$  in a female stable population with fertility rates  $m_x$ , mortality hazard  $\mu_x$  and survival function  $l_x = e^{-\int_0^x \mu_t dt}$  ( $l_0 = 1$ ), as proposed by Goodman, Keyfitz, and Pullum (1974):

$$CS_a = \int_0^a m_x l_{a-x} dx$$

Following Keyfitz and Caswell (2005), if we take  $a \geq \beta$ , the fertility pattern by age is concentrated around  $\kappa$ , and cumulated fertility (gross reproduction rate because of female model) is  $F = \int_{\alpha}^a m_x dx$ , an approximation of  $CS_a$  using Taylor’s theorem until second order around  $\kappa$  would be:

$$CS_a \approx l_{a-\kappa} \int_{\alpha}^a m_x dx + (l_{a-\kappa})' \int_0^a (x - \kappa) m_x dx + (l_{a-\kappa})'' \int_0^a \frac{(x - \kappa)^2}{2} m_x dx \approx F_a l_{a-\kappa} + \frac{\sigma^2}{2} F_a (l_{a-\kappa})'' \approx F_a l_{a-\kappa} [1 + \frac{\sigma^2}{2} \frac{(l_{a-\kappa})''}{l_{a-\kappa}}]$$

The second Taylor’s term is null because  $\int_{\alpha}^a x m_x dx = \kappa F_a$ . The formula tells us that the survival experience depends of cumulated fertility (possible cases) and child survival (success cases), but also how concentrated is the fertility around  $\kappa$  and the survival curvature around.

From the mother perspective, the time in lost is a terrible situation. We call *Mean Time Spent with Lost* (MTSL) the absolute measure of the expected total time with a death daughter that lived a mother aged  $a$ , which can be expressed in terms of a temporary expected lost years measure in line with  $e^{\dagger}()$ :

$$MISL_a = \int_0^a m_x \int_0^{a-x} d_t e_{t,a-x-t} dt dx = \int_\alpha^a m_x e_{0|a-x}^\dagger dx$$

Where  $d_t$  is the death distribution from birth,  $e_{0,a-x-t}$  is the life expectancy at birth until age  $a-x-t$  and  $e_{0|a-x}^\dagger$  a temporary dispersion measure. But most interesting would be to compare these lost years with the expected time that mothers would expect to live with their daughters. We call this the *intensity* of loss (*Intensity Time Lost (ITL)*): a ratio between expected time in “lost” and expected time in “life”, that allows to make comparisons in time and between populations.

$$ITL_a = \frac{\int_\alpha^a m_x e_{a-x}^\dagger dx}{\int_\alpha^a m_x e_{a-x} dx}$$

Considering woman that losted a child, different mean age at child loss ( $MAL$ ) impacts the mother experience. This relation can be derived by starting with the mother age  $x+t$  at each death child age at death  $t$ , weighted by the fertility and survival function. In it,  $\kappa$  is the mean age at childbirth for womens aged  $a$ ,  $MAD_{a-x}$  refers to the mean age at death for newborns that die before  $a-x$ ,  $F_a$  is the cumulated fertility for a women aged  $a$ :

$$MAL_a = \frac{\int_0^a m_x \int_0^{a-x} \frac{d_t(x+t)dt}{d_{0,a-x}}} {\int_0^a m_x} MAL_a = \frac{\int_0^a m_x \int_0^{a-x} \frac{d_t(x+t)dt}{d_{0,a-x}}} {\int_0^a m_x} MAL_a = \frac{\int_0^a m_x x}{F_a} + \frac{\int_0^a m_x MAD_{a-x}}{F_a} MAL_a = \kappa + \frac{\int_0^a m_x MAD_{a-x}}{F_a}$$

This is the sum

We now consider the consequences of a change  $\delta$  in mortality in the range  $[0, a-\alpha]$ , where  $\alpha$  is the start age of fertility risk. A discrete approximate of the effect on  $CS_a$  given an absolute change in mortality, considering that  $m_{x,\delta} = m_x + \delta$  (Wrycza and Baudisch (2012)), and  $l_{a-x} = e^{-\int_0^{a-x} (\mu_t + \delta)}$ :

$$CS_{(a)}^\delta = \int_\alpha^a m_x l_{a-x} e^{-\delta(a-x)} dx$$

We get the derivative of  $dCS_{(a)}^\delta/d\delta$  evaluated near zero (Keyfitz and Caswell (2005)) to find the effects of adding  $\delta$  to the age-specific death rates of daughters:

$$= -a \int_\alpha^a m_x l_{a-x} dx + \int_\alpha^a m_x l_{a-x} x dx$$

Since  $CS_a = \int_\alpha^a m_x l_{a-x} dx$ , we can rewrite equation (??) as

$$\frac{dCS_a^\delta}{d\delta} = -a CS_a + \int_\alpha^a x m_x l_{a-x} dx$$

Dividing both sides by  $CS_a$  in an discrete approzimation we get

$$\frac{\Delta CS_a}{CS_a} \approx -(a - \kappa_{CS}) \Delta \delta$$

In this equation, the expected change in child survival is inversely proportional to the difference between maternal age  $a$  and the mean age of the mother at the birth of her surviving daughters  $\kappa_{CS}$  (always smaller

than  $\kappa$ ). The magnitude of the change depends negatively on the age distribution of the surviving offspring. This is intuitive considering that younger children experience longer periods of exposure to risk.

The relationship is less clear for proportional changes in mortality given the interaction of birth and mortality rates (Keyfitz and Caswell (2005)). Considering that  $m_{x,\delta} = m_x(1+\delta)$  and  $l_{a-x,\delta} = e^{-\int_0^{a-x} \mu_t(1+\delta)dt} = l_{a-x}^{(1+\delta)}$ :

$$CS_a^\delta = \int_\alpha^a m_x l_{a-x}^{(1+\delta)} dx$$

Using the derivative of  $dCS_a^\delta/d\delta = \log(l_{a-x})l_{a-x}^{(1+\delta)}$ , reversing between  $t$  and  $x$ :

$$\frac{dCS_a}{d\delta} = \int_\alpha^a m_x l_{a-x} \log(l_{a-x}) dx = - \int_0^a m_x l_{a-x} \int_0^{a-x} \mu_t dt dx = - \int_0^a \mu_t \int_0^{a-t} m_x l_{a-t-x} dx dt$$

Dividing by  $CS_a$  and multiplying by  $\delta$  get the final expression. This is the negative of cumulative hazard to age  $a$  times a factor  $0 \leq \frac{CS_{a-t}}{CS_a} \leq 1$  (strictly incremental because  $m_x \geq 0$  and  $l_x \geq 0$ ), that will give more weight to mortality in first ages.

$$\frac{\Delta CS_a}{CS_a} = - \left[ \int_0^a \mu_t \frac{CS_{a-t}}{CS_a} dt \right] \Delta \delta$$

```
# age survivor spline

lx_fun <- function(data, age){
  splinefun(data$x, data$lx, method = "monoH.FC")(age)
}

# function to calculate all the measures

CS_results <- function(country, period, a){

  # filter country
  country = LA %>% filter(name == country, Period==period)%>%
    mutate(dx = c(-diff(lx),lx[101]),
           Lx = ifelse(x==0, Lx_n,
                      ifelse(x==100, ex,
                             lx-dx/2)),
           Lx = as.numeric(Lx),
           ex = rev(cumsum(rev(Lx)))/lx)

  # calculate
  CS = sum(unlist(sapply(1:a, function(i) country$asfr[i]*country$lx[a-(i-1)]/100000)), na.rm = T)
  Fa = sum(country$asfr[1:a], na.rm = T)
  CS_prob = CS/Fa
  k = sum(country$asfr[1:a]*0:(a-1), na.rm = T)/Fa
  CS_aprox = Fa * country$lx[a-k+1]/100000
  error_aprox = (CS-CS_aprox)/CS*100

  MTSL = 0
  for(i in 1:a){
    for(j in 1:(a-i)){
      MTSL = sum(MTSL,
                 country$asfr[i] * country$dx[j]/100000 * sum(country$Lx[j:a])/country$lx[j],
```

```

        na.rm=T)
    }
  }
  MTSA = 0
  for(i in 1:a){
    for(j in 1:(a-i)){
      MTSA = sum(MTSA,
                  country$asfr[i] * sum(country$Lx[j:a])/country$lx[j],
                  na.rm=T)
    }
  }
  ITL = MTSL/MTSA * 100
  m_MAD = 0
  for(i in 1:a){
    for(j in 1:(a-i)){
      m_MAD = sum(m_MAD,
                  country$asfr[i] * country$dx[j]/100000 * (j-1),
                  na.rm=T)
    }
  }
  MAL = k + m_MAD/Fa
  k_CS = sum(unlist(sapply(1:a, function(i) country$asfr[i]*country$lx[a-(i-1)]/100000*i)), na.rm = T)/a
  abs_change = -(a-k_CS)

  # Results
  return(data.frame(CS,Fa,CS_prob,k,CS_aprox,error_aprox,MTSL,ITL,MAL))
}

```

## Data

Mortality and fertility data was obtained from the last revision of population prospects by UN (Nations (2019)). We smoothed female  $l_x$  using cubic-splines constrained to monothonic descense, taking later  $L_0$  and  $T_{100}$  as inputs for year-person calculations. For splitting fertility five groups was used quadratic optimization approach by Michalski and Gorlishchev (2018), with an desirable property for our purpose which is a good fitting in parity. Also was assumed an unique female percentage of newborns of 0.49.

```

# Get upp locations
data(UNlocations)

# countries of LA
paises_LA_seleccion <- c(192,214,332,388,630,188,222,320,340,484,558,591,32,68,76,152,170,218,600,604,8)

# get fertility data
data(tfr)
data(percentASFR)
tfr_paises_LA <- tfr %>% select(-last.observed) %>% gather(Period,tfr,-country_code,-name) %>% filter(c

asfr_paises_LA <- percentASFR %>% gather(Period, asfr, -name, -country_code, -age) %>%
  mutate(x = as.numeric(substr(age,1,2)), asfr=asfr/100) %>%
  left_join(tfr_paises_LA, by=c("country_code","name","Period")) %>%
  mutate(asfr = asfr/5 * tfr) %>%
  select(-age) %>%
  filter(country_code %in% paises_LA_seleccion)

```

```

# lt must downloaded manually
lt_paises_LA <- readxl::read_xlsx("WPP2019_MORT_F17_3_ABRIDGED_LIFE_TABLE_FEMALE.xlsx", sheet = 1, range = "A1:Z1000")



x = "Age (x)",
mx = "Central death rate m(x,n)",
qx = "Probability of dying q(x,n)",
dx = "Number of deaths d(x,n)",
lx = "Number of survivors l(x)",
Lx_n = "Number of person-years lived L(x,n)",
ex = "Expectation of life e(x)" %>%

select(country_code, Period, x, mx, qx, dx, lx, Lx_n, ex) %>%
filter(country_code %in% paises_LA_seleccion)

# smooth both
asfr_paises_LA_s <- data.frame(country_code=0, Period = " ", x = 0, asfr = 0, stringsAsFactors = F)
lx_paises_LA_s <- data.frame(country_code=0, Period = " ", x = 0, lx = 0, stringsAsFactors = F)
countries = unique(asfr_paises_LA$country_code)
for(c in 1:length(countries)){
  for(p in unique(asfr_paises_LA$Period)){
    cp = asfr_paises_LA %>% filter(country_code==countries[c], Period==p) %>% select(asfr)
    if(any(is.na(cp$asfr))){next}
    out <- tryCatch(QOSplit(c(0, cp$asfr), L=seq(10, 45, 5), AgeInt=rep(5, 8)),
                     error = data.frame(Age=10:49, ASFR=rep(NA, 40)))
    asfr_paises_LA_s <- rbind(asfr_paises_LA_s, data.frame(country_code = countries[c], Period = p, x =
                                                         out$Age, asfr = out$ASFR))
    cp <- lt_paises_LA %>% filter(country_code==countries[c], Period==p) %>% select(x, lx)
    out <- lx_fun(cp, age=0:100)
    lx_paises_LA_s <- rbind(lx_paises_LA_s, data.frame(country_code=countries[c], Period = p, x = 0:100))
  }
}

# join everything
LA <- lx_paises_LA_s %>% slice(-1) %>% mutate(country_code=as.integer(country_code)) %>%
  left_join(asfr_paises_LA_s %>% slice(-1), by=c("country_code", "Period", "x")) %>%
  left_join(tfr_paises_LA, by=c("country_code", "Period")) %>%
  left_join(lt_paises_LA %>%
            select(country_code, Period, x, Lx_n, ex),
            by = c("country_code", "Period", "x")) %>%
  mutate(q = (1-lx/100000),
         name = as.character(name),
         tfr = tfr*im,
         asfr = asfr*im)
LA$name[LA$name == "Bolivia (Plurinational State of)"] = "Bolivia"

```

An negative relation between lity until 10 and TFR is shown in [@ref\(fig:plot\\_tfr\\_q010\)](#). Specific paths like Argentina and Guatemala ows very different demographic profiles during second part of XX century.

```

ggplot(LA %>% filter(x==10), aes(x=tfr, y=q))+
  geom_point(aes(color=Period)) +
  geom_smooth(se = F, color = "grey") +
  geom_line(data = LA %>% filter(x==10, name %in% c("Guatemala")),
            aes(group=name), colour=2, size = 1.5) +
  geom_line(data = LA %>% filter(x==10, name %in% c("Argentina")),
            aes(group=name), colour=4, size = 1.5) +
  theme_classic() + labs(x="GRR", y="q(0,10)")

```

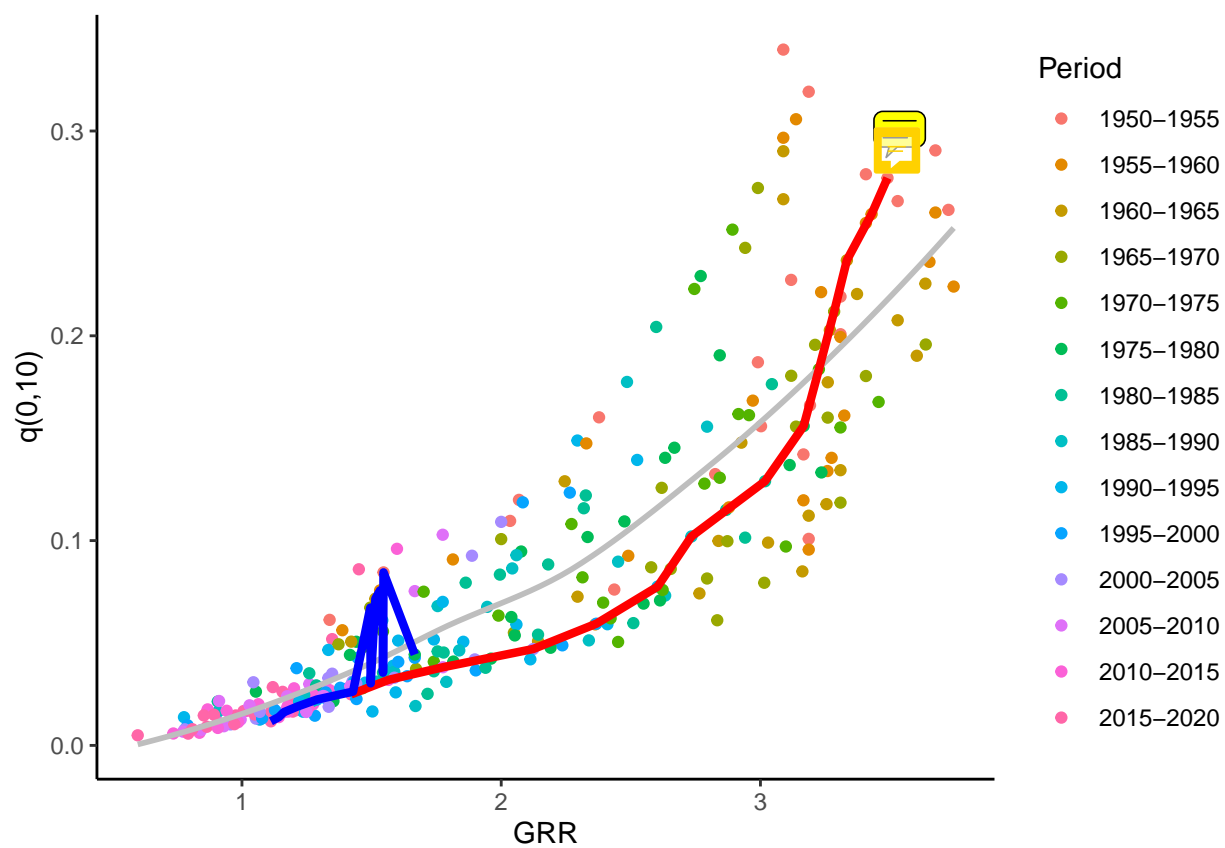


Figure 1: Female probability of death in first 10 years old by gross reproduction rate. Latin America countries in period 1950-2015

## Application

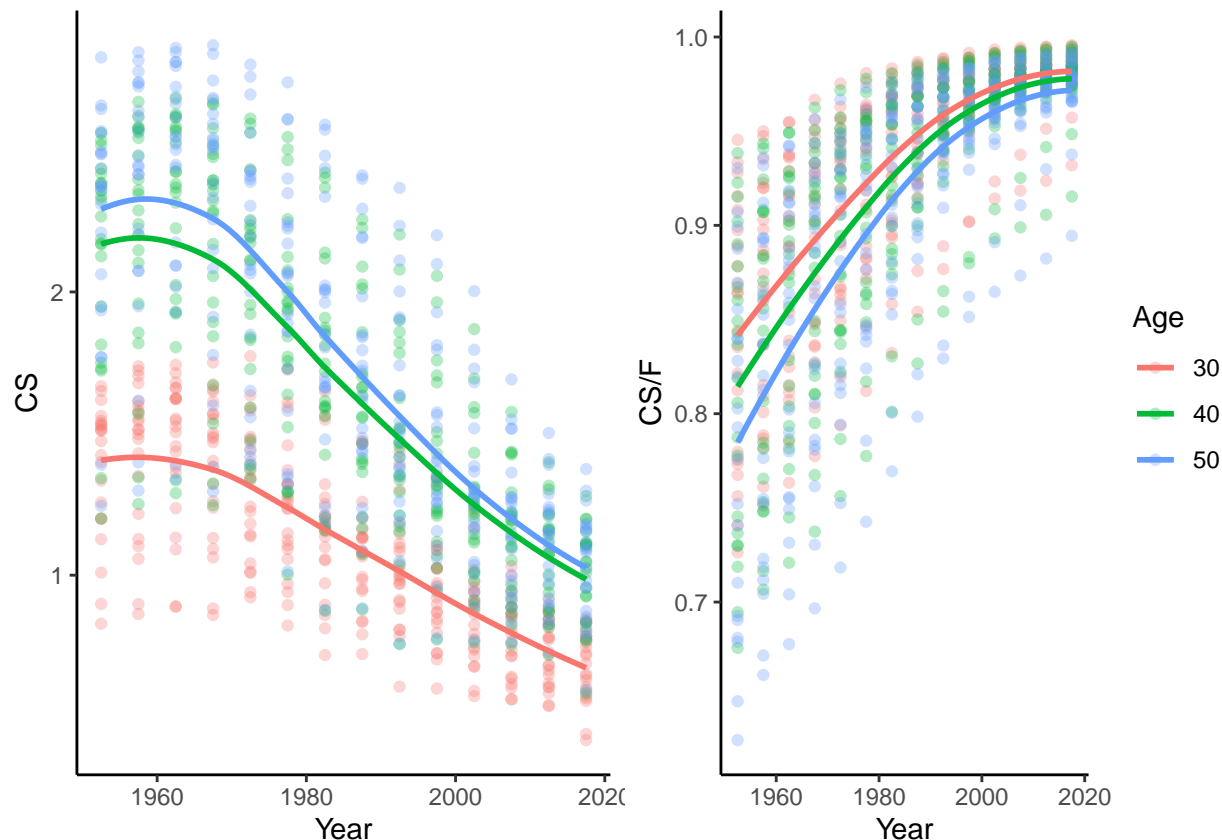
```
CS_outs <- as.data.frame(expand.grid(name = unique(LA$name),
                                     Age = c(30,40,50),
                                     Period = unique(LA$Period)))

for(i in 1:nrow(CS_outs)){
  CS_outs[i,4:12] = CS_results(country = CS_outs[i,1],
                              period = CS_outs[i,3],
                              a = CS_outs[i,2])}

CS_outs <- CS_outs %>% mutate(Year = as.integer(substr(Period, 1, 4)) + 2.5)
```

Between 1950-1955 and 2010-2015 the portion of surviving daughters  $\frac{CS}{F_a}$  at age 50 passed from a simple mean of 0.78 to 0.97, also with a decrease of 1.14 in the expected descendants at same age. Although this is not a real measure but a theoretical one if the rates would have been kept constant during enough time.

```
a <- ggplot(CS_outs, aes(x=Year, y=CS, color=factor(Age)))+
  geom_point(alpha = .3) +
  geom_smooth(se = F) +
  theme_classic() +
  labs(y = "CS") + theme(legend.position = "none")
b <- ggplot(CS_outs, aes(x=Year, y=CS_prob, color=factor(Age)))+
  geom_point(alpha = .3) +
  geom_smooth(se = F) +
  theme_classic() +
  labs(y = "CS/F") +
  scale_colour_discrete(name="Age")
grid.arrange(a,b,ncol=2)
```



The approximation in  $()$  is very accurate, improving with the years. In the case of Guatemala, is attributable to the  $l_x$  rectangularization process.

```
a <- ggplot(CS_outs, aes(x=CS, y=error_aprox, color=Year))+
  geom_point(alpha = .6) +
  theme_classic() +
  theme(legend.position = "none") +
  labs(y = "%")
b <- LA %>% filter(name=="Guatemala") %>% ggplot() +
  geom_line(aes(x=x, y=lx/100000, color=Period)) +
  theme_classic() +
  theme(legend.position = "none") +
  labs(y = "")
c <- LA %>% filter(name=="Guatemala", between(x, 15,50)) %>% ggplot() +
  geom_line(aes(x=x, y=asfr/tfr*100, color=Period)) +
  theme_classic() +
  theme(legend.position = "none") +
  labs(y = "%")
grid.arrange(a, grid.arrange(b,c,ncol=1), ncol=2)
```

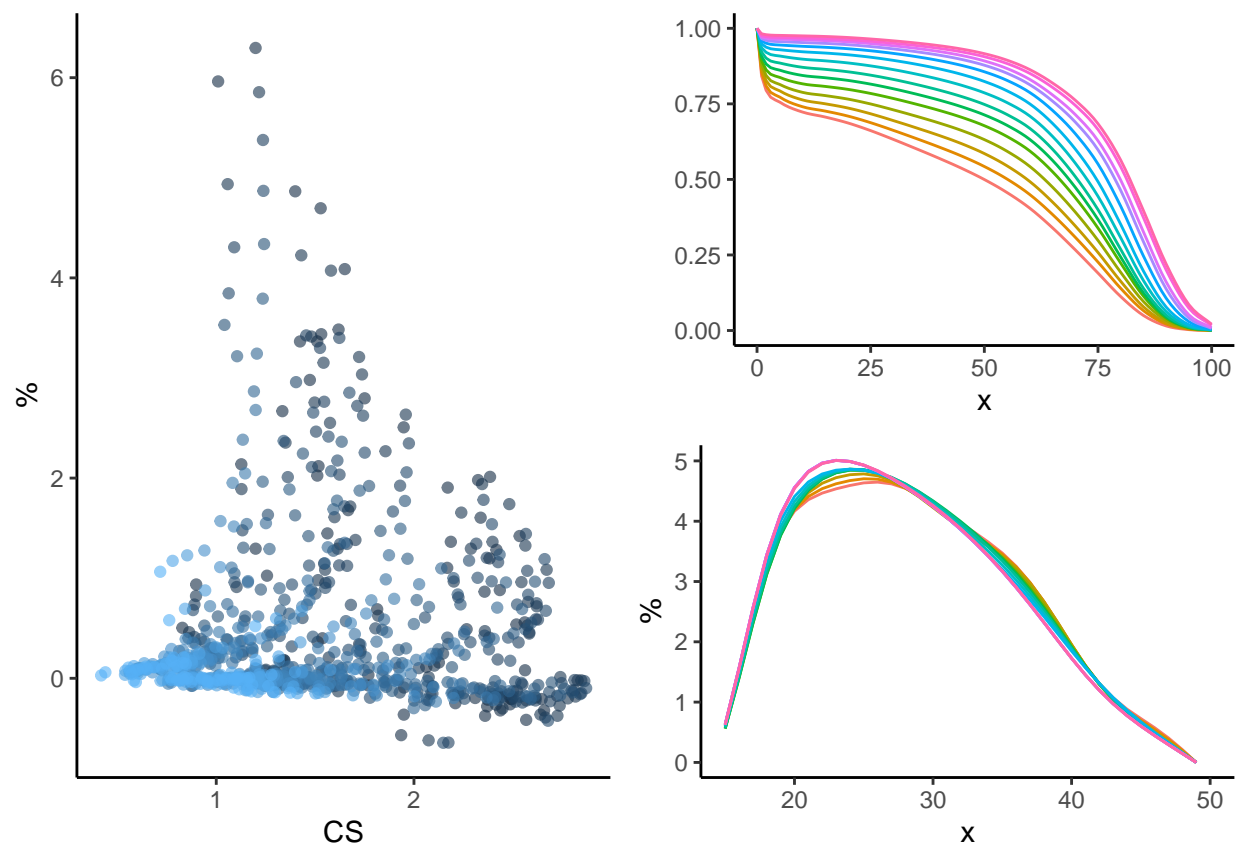


Figure 2: Error (%) in approximation, 1950 (darker) to 2015 (lighter) (left). Guatemala case (right)

The expected time in years that a mother aged 30 passed with a death son was around 4% in some countries at middle XX Century. When increasing age, the survival experience depends less on infant mortality, and the distribution is around less than 2% for womans aged 50, converging to 0 on time.



```
ggplot(CS_outs)+ geom_point(aes(x=Fa, y=ITL, color=Year), alpha = .3) +
  theme_classic()+
  theme(legend.position = "none") +
  labs(y = "%")+
  facet_grid(~Age)
```

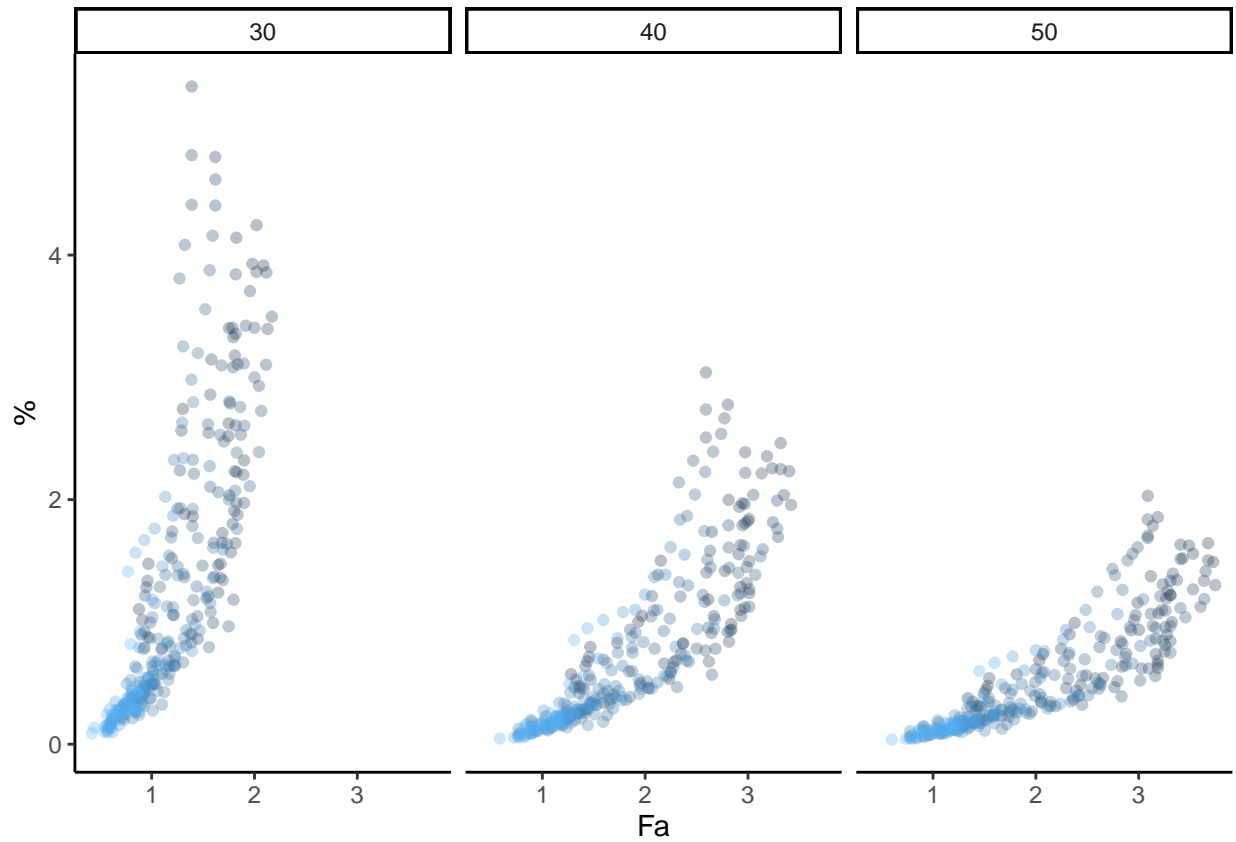
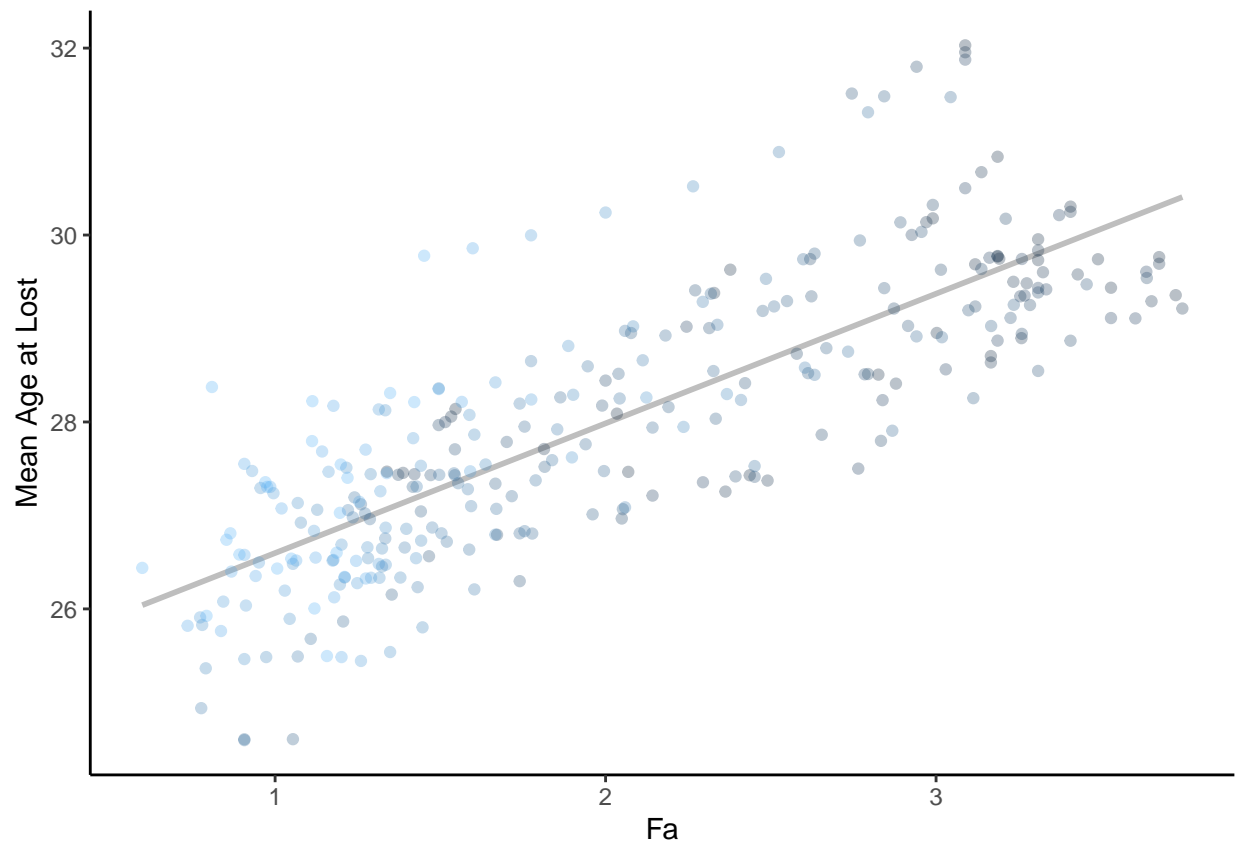


Figure 3: Intensity Time Lost of womans aged 30, 40 and 50. Years 1950 (darker) to 2015 (lighter)

For those woman that lost a daughter, more fertility more the mean age.

```
ggplot(CS_outs %>% filter(Age==50))+
  geom_smooth(aes(x=Fa, y=MAL), method = "lm", col="grey", se = F)+
  geom_point(aes(x=Fa, y=MAL, color=Year), alpha = .3) +
  theme_classic()+
  theme(legend.position = "none") +
  labs(y = "Mean Age at Lost")
```



## Discussion



## pendings

Here an independent assumption is done in the relationship between fertility and descendants survival within population. Both assumptions will be considered at the end.



In terms of mortality, this measure is much more related to first ages (child and youth) than adult or adult ages, so we can model a H-P pattern and derive some of their parameters...

We can take cohort rates to compare between countries, relaxing the assumptions in stable model...



It is known that fertility is influenced by child mortality, and that there is a heterogeneous behaviour within. We can add fractional variables to take account of that and express correlation between subgroups behaviour...



## 5. Bibliography

Bongaarts, John. 1987. "The Projection of Family Composition over the Life Course with Family Status Life Tables." In *Family Demography: Methods and Their Application*, edited by John Bongaarts and Thomas Burch. Oxford: IUSSP Series, Oxford University Press.

Brass, W. 1953. "The Derivation of Fertility and Reproduction Rates from Restricted Data on Reproductive Histories." *Population Studies* 7 (2): 137. <https://doi.org/10.2307/2172029>.

Goodman, Leo A, Nathan Keyfitz, and Thomas W Pullum. 1974. "Family Formation and the Frequency of Various Kinship Relationships." *Theoretical Population Biology* 5 (1): 1-27.

Keyfitz, Nathan, and Hal Caswell. 2005. *Applied Mathematical Demography*. Vol. 47. Springer.

- Lotka, Alfred J. 1931. "Orphanhood in Relation to Demographic Factors." *Metron* 9 (2): 37–109.
- Michalski, Grigoriev, Anatoli I, and Vasily P and Gorlishchev. 2018. "R Programs for Splitting Abridged Fertility Data into a Fine Grid of Ages Using the Quadratic Optimization Method." Max Planck Institute for Demographic Research, Rostock, Germany.
- Nations, United. 2019. *World Population Prospects 2019: Data Booklet*. <https://doi.org/https://doi.org/https://doi.org/10.18356/3e9d869f-en>.
- Wrycza, Tomasz, and Annette Baudisch. 2012. "How Life Expectancy Varies with Perturbations in Age-Specific Mortality." *Demographic Research* 27: 365–76.