Manipulating the R parameter in the QS fertility model

Carl Schmertmann

01 Nov 2019

Purpose

A quick illustration of how changing the R parameter affects the QS fertility model described in

CP Schmertmann, 2003. *A System of Model Fertility Schedules with Graphically Intuitive Parameters*. **Demographic Research** Vol. 9-5 (10 Oct 2003). https://dx.doi.org/10.4054/DemRes.2003.9.5

The main idea is to find a continuous function (specifically, a restricted quadratic spline) that approximates the observed f_x values. A QS curve has four parameters:

1. R : the peak level of fertility
2. α : the lowest age at which fertility is positive
3. P : the age at which fertility reaches its peak value R
4. H : the age $> P$ at which fertility falls to half its peak value, $R/2$.

Define the $QS(\,)$ function

x is the vector of ages at which to calculate interpolated rates

```
QS = function(x, R=.20,alpha=15,P=28,H=34) {

  w = min(.75, .25+.025*(P-alpha))
  beta = max (  min(50, 4*H-3*P) , (4*H-P)/3 )

  D = P-20           # delay index
  C = (P+50)/2 - H  # control index

  knot    = vector("numeric",5)
  knot[1] = alpha
  knot[2] = alpha + w*(P-alpha)
  knot[3] = P
  knot[4] = (P+H)/2
  knot[5] = (H+beta)/2

  A      = matrix(NA,5,5)
  target = vector("numeric",5)

  # target value at P=1
  A[1,]     = (pmax(P-knot, 0))^2
  target[1] = 1

  # target value at H=1/2
  A[2,]     = (pmax(H-knot, 0))^2
  target[2] = 1/2

  # target value at beta=0
  A[3,]     = (pmax(beta-knot, 0))^2
  target[3] = 0
```

```
  # target slope at P=0
  A[4,]      = 2*pmax(P-knot, 0)
  target[4] = 0

  # target slope at beta=0
  A[5,]      = 2*pmax(beta-knot, 0)
  target[5] = 0

  # calculate thetas
  theta = solve(A,target)

  tmp = (pmax( outer(x,knot,"-") , 0))^2

  y = (x >= alpha) * (x <= beta) * R * (tmp %*% theta)  # ASFRs

  return(list(R=R,alpha=alpha,P=P,H=H,
              theta=theta,knot=knot,w=w,beta=beta,delay=D,control=C,
              x=as.vector(x) ,y=as.vector(y) ))

} #QS
```

The $\phi(x)$ function

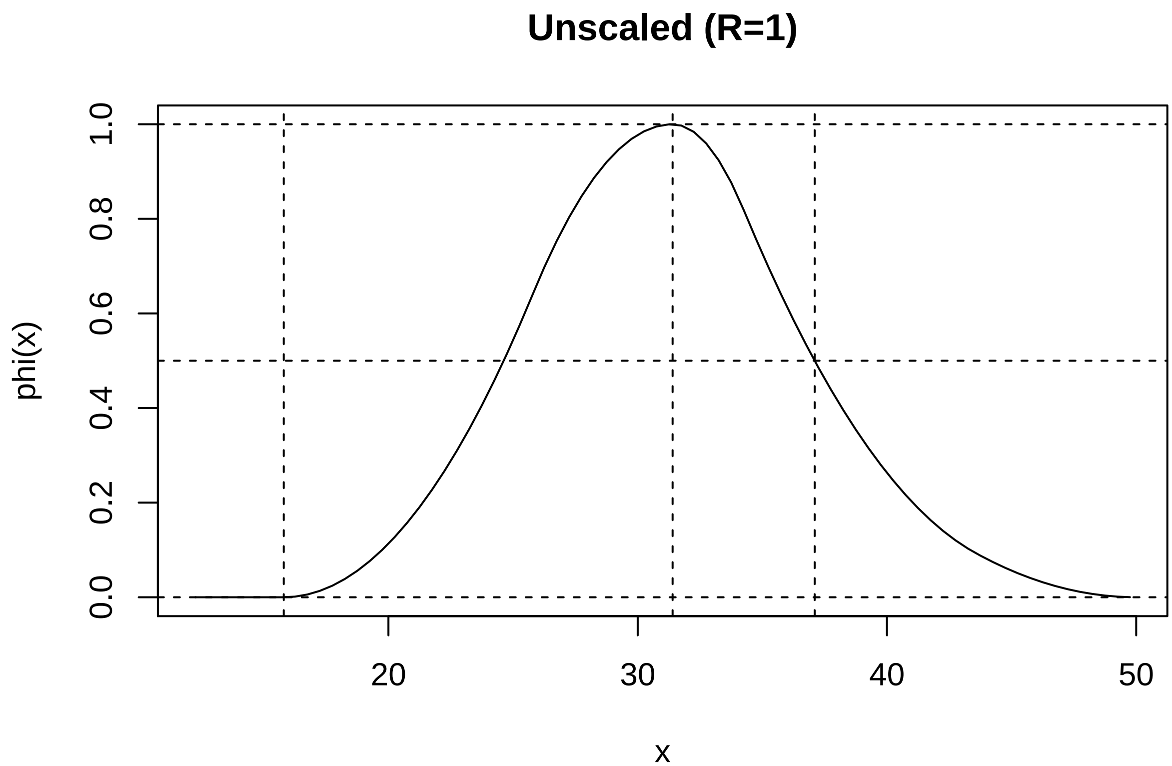
The shape of a fitted QS schedule depends on the age parameters – α, P, and H. In this example I'll use

$$(\alpha, P, H) \;=\; (15.8, 31.4, 37.1)$$

which matches the shape for 2016 Swedish fertility rates in the Human Fertility Database.

First produce a QS schedule over a fine grid of ages $12.25, 12.75, \ldots, 49.25, 49.75$.

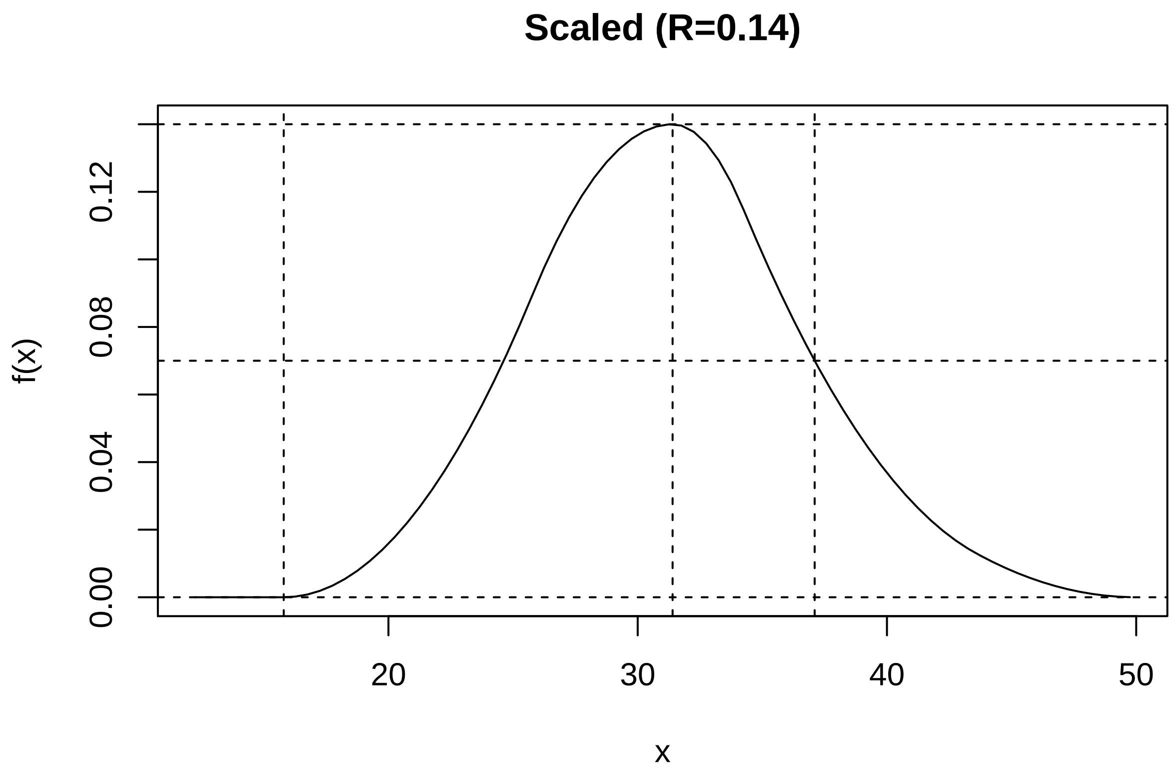
```
x = seq(from=12.25, to=49.75, by=0.5)

asfr = QS(x, alpha=15.8, R=1, P=31.4, H=37.1)

plot(asfr, type='l', xlab='x', ylab='phi(x)', main='Unscaled (R=1)')
abline(v=c(15.8, 31.4, 37.1), h=c(0, 0.5, 1), lty='dashed')
```



Direct use of R: Define the maximum fertility rate

If we change R from 1 to some other value (e.g., $R = 0.14$), we are changing the fertility schedule from $f(x) = 1 \cdot \phi(x)$ to $f(x) = R \cdot \phi(x)$, and R is the new height at the mode:

```
asfr = QS(x, alpha=15.8, R=0.14, P=31.4, H=37.1)

plot(asfr, type='l', xlab='x', ylab='f(x)', main='Scaled (R=0.14)')
abline(v=c(15.8, 31.4, 37.1), h=0.14*c(0, 0.5, 1), lty='dashed')
```



Indirect use of R: Match a target TFR value

We could also choose the value of R to ensure that $TFR = \int_{\alpha}^{50} f(x)dx$ matches some specific value. This value would be

$$R^* = \frac{TFR}{\int_{\alpha}^{50} \phi(x)dx}$$

For example, if we wanted a schedule with

$$TFR = 1.30,\ \alpha = 15.8,\ P = 31.4,\ H = 37.1$$

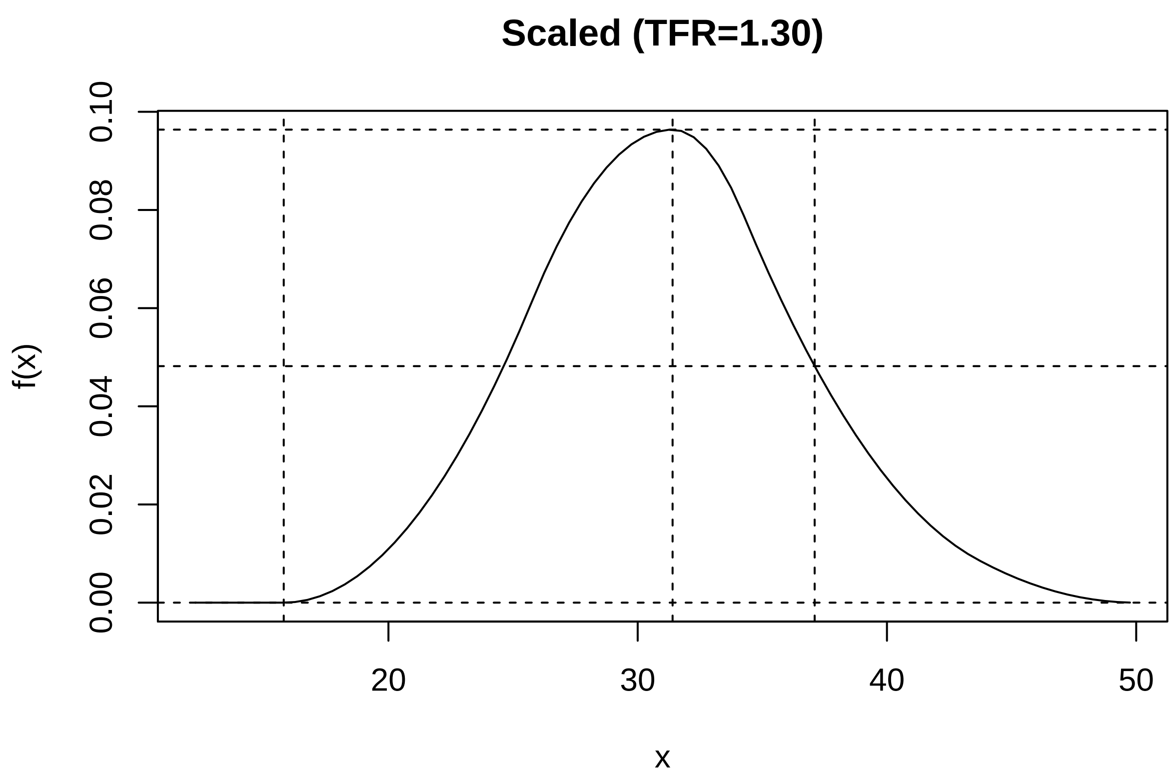
we could do the following

```
phi = QS(x, alpha=15.8, R=1, P=31.4, H=37.1)

# 0.50 because that's the width of intervals on the age grid
integral_phi = sum(phi$y) * .50

print(integral_phi)
```

```
## [1] 13.48997
```

```
Rstar = 1.30 / integral_phi

asfr = phi = QS(x, alpha=15.8, R=Rstar, P=31.4, H=37.1)
```

```r
plot(asfr, type='l', xlab='x', ylab='f(x)', main='Scaled (TFR=1.30)')
abline(v=c(15.8, 31.4, 37.1), h=Rstar*c(0, 0.5, 1), lty='dashed')
```




```r
## check
TFR = sum(asfr$y) * .50
print(TFR)
```

```
## [1] 1.3
```

Other indirect uses of R

We could also get creative, and choose R to match other features of the ASFR schedule that depend on the scaling, like

- average number of children born after age 30
- probability of having 2 children by age 35 (if we assume that births are Poisson distributed…)
- etc.