kyotil Package Vignette

Youyi Fong

August 22, 2020

1 Computing marginal risk using the marginal.risk function

To plot the disease risk as a function of a marker (Gilbert et al., 2014), we use the following formula:

$$\begin{aligned}
&\Pr(Y=1|s) \\
&= \int \Pr(Y=1|s,z)\, f(z|s)\, dz \\
&= \int \Pr(Y=1|s,z)\, \frac{f(s|z)\, f(z)}{f(s)} dz \\
&= \int \Pr(Y=1|s,z)\, \frac{f(s|z)\, f(z)}{\int f(s|z)\, f(z)\, dz} dz \\
&= \frac{\sum_i \Pr(Y=1|s,z_i)\, f(s|z_i)}{\sum_i f(s|z_i)}
\end{aligned}$$

where $f(s|z)$, the density of s conditional on z, can be estimated by a fitting a linear regression model of s on z using the data. If the data is collected according to a two phase sampling design and z_i has an inversion probability sampling weight w_i, then the formula can be updated to:

$$\Pr(Y=1|s) = \frac{\sum_i w_i \Pr(Y=1|s,z_i)\, f(s|z_i)}{\sum_i w_i f(s|z_i)},$$

and the estimation of the conditional distribution $f(s|z)$ can also include the weights to improve efficiency.
For more info on how to use the function, its help page has an example.

References

Gilbert, P., Gabriel, E., Miao, X., Li, X., Su, S.C., Parrino, J. et al (2014), "Fold-rise in antibody titers by gpELISA is an excellent correlate of protection for a herpes zoster vaccine, demonstrated via the vaccine efficacy curve," *Journal of Infectious Diseases*, 210, 1573–81.