Background Knowledge for EPI511

Review of standard epi terminology

Measures of Disease Occurrence

Incidence, Prevalence, Risk, Relative Risk

Observational Study Types

Cohort, Case-Control, Matched Case-Control, Cross Sectional, Nested Case Control, Case-Cohort, Ecological

Confounder (but for real whats a confounder?)

Rothman and Greenland have a pretty good definition
1. Risk factor for the response
2. Associated with the exposure under study in the source population
3. Not affected by the exposure or the response (like cant be intermediate step in the causal path)

Risk, Rates, and some Maths

Hazard Function

Let T be the survival time for an individual in a cohort
Let $P = h \times m$ where
- P is the observation time of the study (5 years)
- h is an interval of time (6 months or .5 years)
- m will therefore be the number of intervals (10 in this case)

Intervals can now be denoted as $[t_i, t_{i+1})$ where
$t_i = (i-1) \times P/m$

The probability of dying within a particular interval can be written as
$\pi(t_i) = Pr(t_i \leq T < t_{i+1} | T \geq t_i)$
which is basically saying the probability of dying in interval $[t_i, t_{i+1})$ given you made it to t_i

This can be approximated by
$\approx \lambda(t_i) \times h$
where $\lambda(t_i)$ is the hazard function or the **instantaneous** probability of failure

Remember that the hazard function is a rate which is shown by rearranging the above equation
$\lambda(t_i) \approx \pi(t_i)/h$

```
lambda <- pi / h
```

```
lambda = float(pi) / float(h)
```