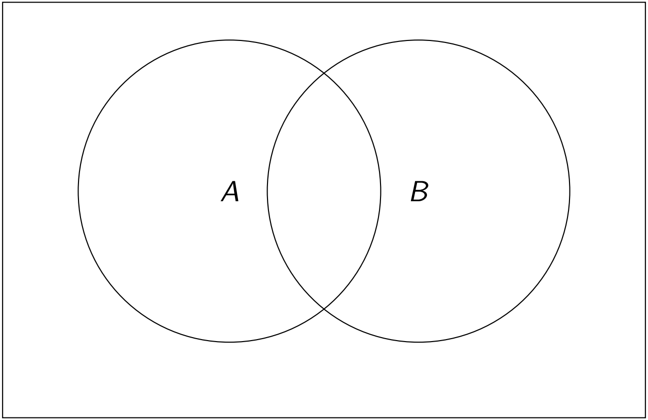
Odds Ratios

Randy Johnson

3/9/2017

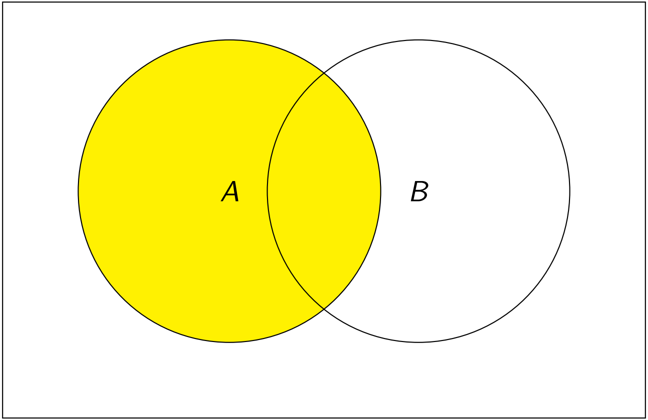
Quick review of probability

This Venn diagram represents the event space for two events, A and B. The area inside of the rectangle is 1.



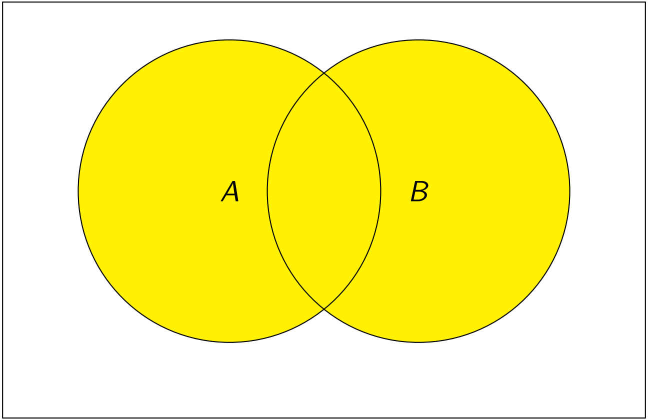
Quick review of probability

This represents $P(A)$.



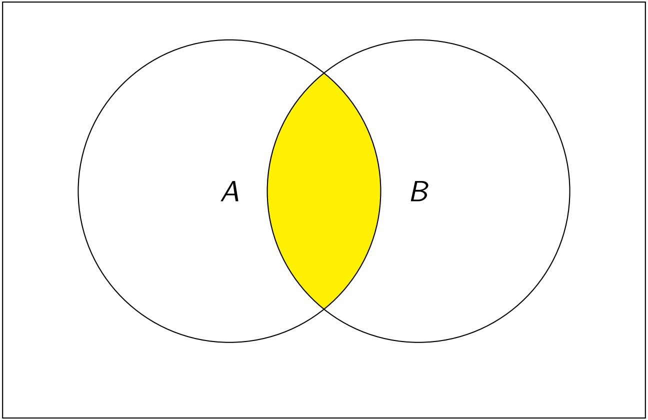
Quick review of probability: Union

This represents $\mathrm{P}(A \cup B)$.



Quick review of probability: Intersection

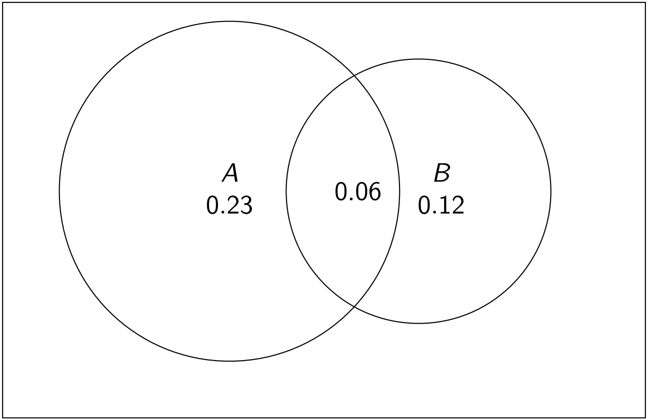
This represents $P(A \cap B)$.



Quick review of probability: Practice

$$P(!(A \cup B)) = 0.59$$
$$P(A) = 0.29$$
$$P(A \cap !B) = 0.23$$

$$P(A \cap B) = 0.06$$
$$P(A|B) = \frac{0.06}{0.18} = 0.33$$
$$P(A|!B) = \frac{0.23}{0.82} = 0.28$$



Definition: Odds of D given E

Given a disease outcome, D, an exposure, E, and

	$D = 1$	$D = 0$	row total
$E = 1$	$n_{1,1}$	$n_{1,0}$	$n_{1,\bullet}$
$E = 0$	$n_{0,1}$	$n_{0,0}$	$n_{0,\bullet}$
col total	$n_{\bullet,1}$	$n_{\bullet,0}$	N

the odds of event D given exposure are

$$\begin{aligned}
\mathrm{P}(D = 1|E = 1) \\
&= \frac{\mathrm{P}(D = 1 \cap E = 1)}{\mathrm{P}(E = 1)} \\
&= \frac{n_{1,1}}{\cancel{N}} \Big/ \frac{n_{1,\bullet}}{\cancel{N}} \\
&= \frac{n_{1,1}}{n_{1,\bullet}}
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}(D = 0|E = 1) \\
&= \frac{\mathrm{P}(D = 0 \cap E = 1)}{\mathrm{P}(E = 1)} \\
&= \frac{n_{1,0}}{\cancel{N}} \Big/ \frac{n_{1,\bullet}}{\cancel{N}} \\
&= \frac{n_{1,0}}{n_{1,\bullet}}
\end{aligned}$$

Definition: Odds of D given E

Given a disease outcome, D, an exposure, E, and

	$D=1$	$D=0$	row total
$E=1$	$n_{1,1}$	$n_{1,0}$	$n_{1,\bullet}$
$E=0$	$n_{0,1}$	$n_{0,0}$	$n_{0,\bullet}$
col total	$n_{\bullet,1}$	$n_{\bullet,0}$	N

the odds of event D given exposure are

$$
\begin{aligned}
odds(D=1|E=1) &= \frac{\mathrm{P}(D=1|E=1)}{\mathrm{P}(D=0|E=1)} \\
&= \frac{n_{1,1}}{\cancel{n_{1,\bullet}}} \Big/ \frac{n_{1,0}}{\cancel{n_{1,\bullet}}} \\
&= \frac{n_{1,1}}{n_{1,0}}
\end{aligned}
$$

Definition: Odds of D given $!E$

Given a disease outcome, D, an exposure, E, and

	$D=1$	$D=0$	row total
$E=1$	$n_{1,1}$	$n_{1,0}$	$n_{1,\bullet}$
$E=0$	$n_{0,1}$	$n_{0,0}$	$n_{0,\bullet}$
col total	$n_{\bullet,1}$	$n_{\bullet,0}$	N

the odds of event D given *no* exposure are

$$\begin{aligned}
&\mathrm{P}(D=1|E=1)\\
&\quad= \frac{\mathrm{P}(D=1\cap E=0)}{\mathrm{P}(E=0)}\\
&\quad= \frac{n_{0,1}}{\cancel{N}} \Big/ \frac{n_{0,\bullet}}{\cancel{N}}\\
&\quad= \frac{n_{0,1}}{n_{0,\bullet}}
\end{aligned}$$

$$\begin{aligned}
&\mathrm{P}(D=0|E=1)\\
&\quad= \frac{\mathrm{P}(D=0\cap E=1)}{\mathrm{P}(E=1)}\\
&\quad= \frac{n_{0,0}}{\cancel{N}} \Big/ \frac{n_{0,\bullet}}{\cancel{N}}\\
&\quad= \frac{n_{0,0}}{n_{0,\bullet}}
\end{aligned}$$

Definition: Odds of D given $!E$

Given a disease outcome, D, an exposure, E, and

	$D=1$	$D=0$	row total
$E=1$	$n_{1,1}$	$n_{1,0}$	$n_{1,\bullet}$
$E=0$	$n_{0,1}$	$n_{0,0}$	$n_{0,\bullet}$
col total	$n_{\bullet,1}$	$n_{\bullet,0}$	N

the odds of event D given *no* exposure are

$$\begin{aligned} odds(D=1|E=1) &= \frac{\mathrm{P}(D=1|E=0)}{\mathrm{P}(D=0|E=0)} \\ &= \frac{n_{0,1}}{\cancel{n_{0,\bullet}}} \bigg/ \frac{n_{0,0}}{\cancel{n_{0,\bullet}}} \\ &= \frac{n_{0,1}}{n_{0,0}} \end{aligned}$$

Definition: Odds Ratio

Given a disease outcome, D, an exposure, E, and

	$D=1$	$D=0$	row total
$E=1$	$n_{1,1}$	$n_{1,0}$	$n_{1,\bullet}$
$E=0$	$n_{0,1}$	$n_{0,0}$	$n_{0,\bullet}$
col total	$n_{\bullet,1}$	$n_{\bullet,0}$	N

the ratio of the odds of disease comparing exposure vs no exposure is

$$
\begin{aligned}
OR(D|E) &= \frac{odds(D=1|E=1)}{odds(D=1|E=0)} \\
&= \frac{n_{1,1}}{n_{1,0}} \Big/ \frac{n_{0,1}}{n_{0,0}} \\
&= \frac{n_{1,1}n_{0,0}}{n_{1,0}n_{0,1}}.
\end{aligned}
$$

Definition: $se(\log OR)$

Given a disease outcome, D, an exposure, E, and

	$D=1$	$D=0$	row total
$E=1$	$n_{1,1}$	$n_{1,0}$	$n_{1,\bullet}$
$E=0$	$n_{0,1}$	$n_{0,0}$	$n_{0,\bullet}$
col total	$n_{\bullet,1}$	$n_{\bullet,0}$	N

the log OR is normally distributed, and its standard error is

$$se(\log OR) = \sqrt{\frac{1}{n_{1,1}} + \frac{1}{n_{1,0}} + \frac{1}{n_{0,1}} + \frac{1}{n_{0,0}}}.$$

Putting everything together

Given a disease outcome, D, and exposure, E, and

	$D = 1$	$D = 0$	row total
$E = 1$	36	54	*90*
$E = 0$	32	89	*121*
col total	*68*	*143*	*211*

what is the OR and 95% CI for the exposure, E? Is this statistically significant?

$$
\begin{aligned}
OR(D|E) &= \frac{36 * 89}{32 * 54} \\
&= 1.85 \\
\log OR &= 0.617 \\
se(\log OR) &= \sqrt{\frac{1}{36} + \frac{1}{54} + \frac{1}{32} + \frac{1}{89}} \\
&= 0.298
\end{aligned}
$$

Putting everything together

Given a disease outcome, D, and exposure, E, and

	$D = 1$	$D = 0$	row total
$E = 1$	36	54	*90*
$E = 0$	32	89	*121*
col total	*68*	*143*	*211*

what is the OR and 95% CI for the exposure, E? Is this statistically significant?

$$\begin{aligned} OR &= 1.85 \\ 95\%\ CI(\log OR) &= 0.617 \pm 1.96 * 0.298 \\ &= (0.031, 1.199) \\ 95\%\ CI(OR) &= (1.03, 3.32) \end{aligned}$$

Logistic Regression: Review

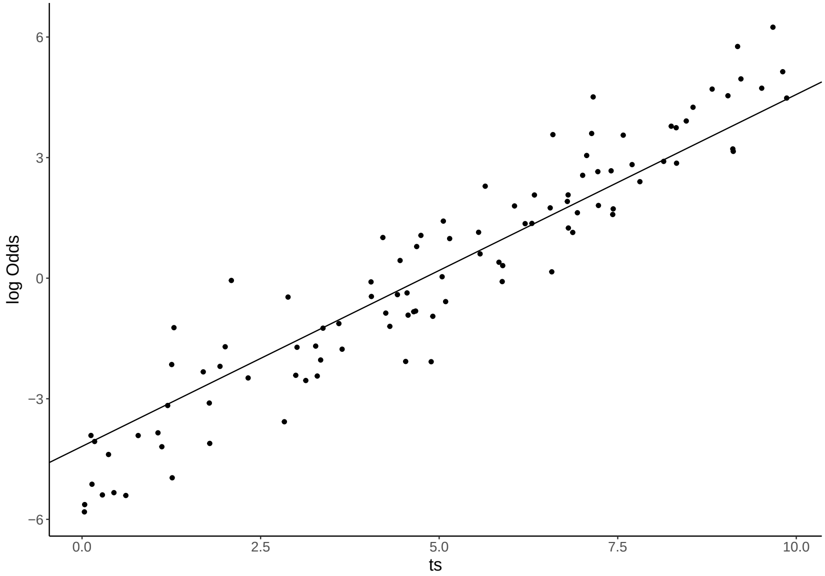
```
set.seed(934875)
n <- 100
dat <- data_frame(ts = runif(100, 0, 10), # tumor size
                  lodds = rnorm(n, (ts - 5)), # log odds of metastisis
                  odds = exp(lodds),
                  p = odds / (1 + odds),
                  metastasis = rbinom(n, 1, p))

model <- glm(metastasis ~ ts, data = dat, family = binomial)

lodds.g <- ggplot(dat, aes(ts, lodds)) +
           geom_point() +
           ylab("log Odds") +
           geom_abline(intercept = model[[1]][1], slope = model[[1]][2])

set.seed(29384)
p.g <- ggplot(dat, aes(ts, metastasis)) +
       geom_jitter(width = 0, height = .05) +
       geom_smooth(method = 'loess', se = FALSE, linetype = 2) +
       geom_line(data = arrange(augment(model), .fitted),
                 aes(ts, exp(.fitted) / (1+exp(.fitted))))
```

Logistic Regression: Review



Logistic Regression: Review