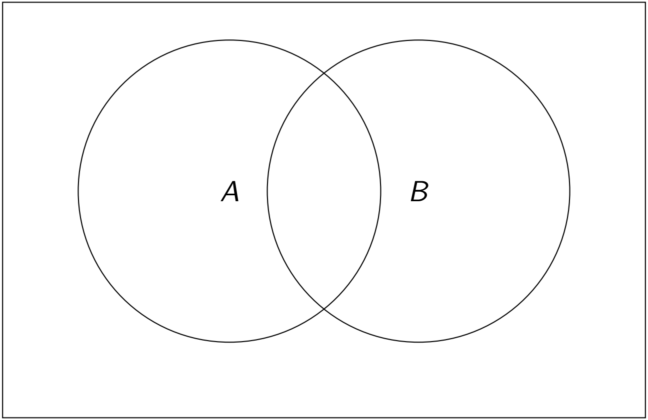
Odds Ratios

Randy Johnson

3/9/2017

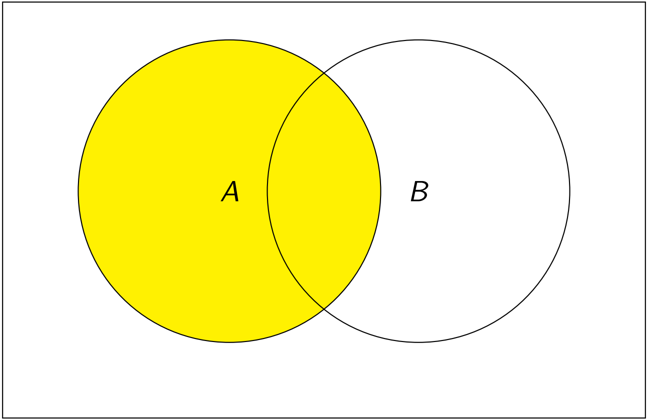
Quick review of probability

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



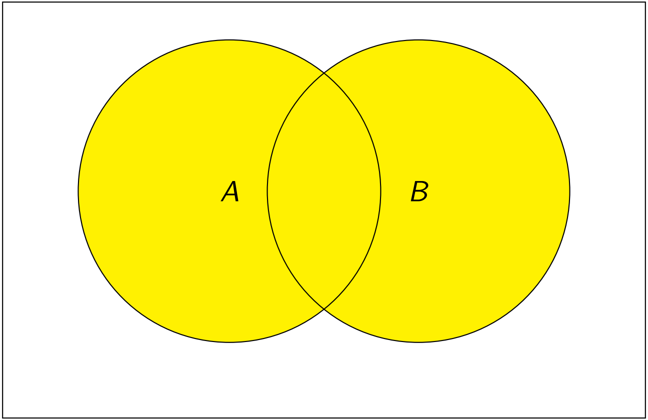
Quick review of probability

This represents $P(A)$.



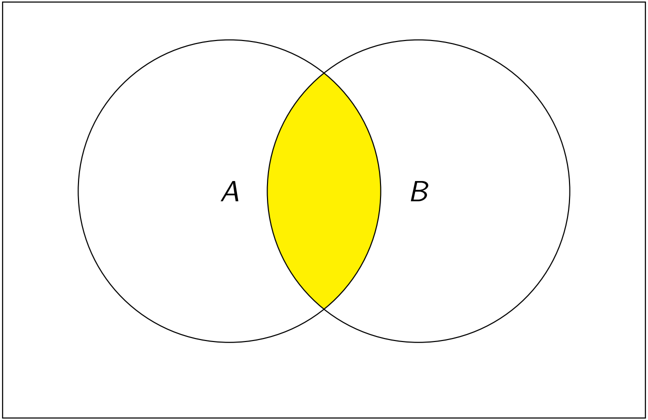
Quick review of probability: Union

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



Quick review of probability: Intersection

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

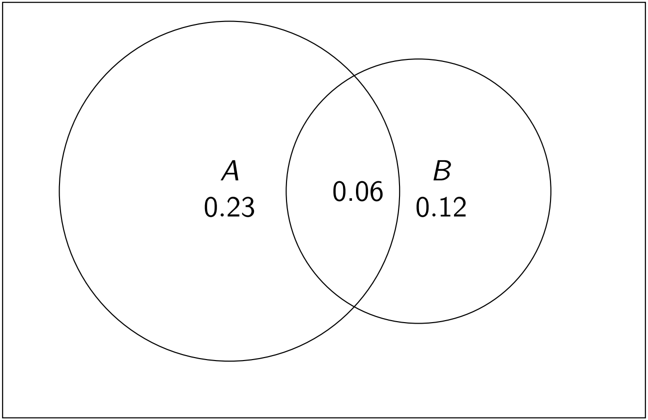


Quick review of probability: Practice

$$\mathrm{P}\left(!\left(A \cup B\right)\right) = \qquad \mathrm{P}(A \cap B) =$$

$$\mathrm{P}(A) = \qquad \mathrm{P}(A|B) =$$

$$\mathrm{P}(A \cap !B) = \qquad \mathrm{P}(A|!B) =$$



Definition: Odds

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 / no exposure are

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

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

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

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

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	
$E=0$	32	89	
col total			

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

Logistic Regression: Review

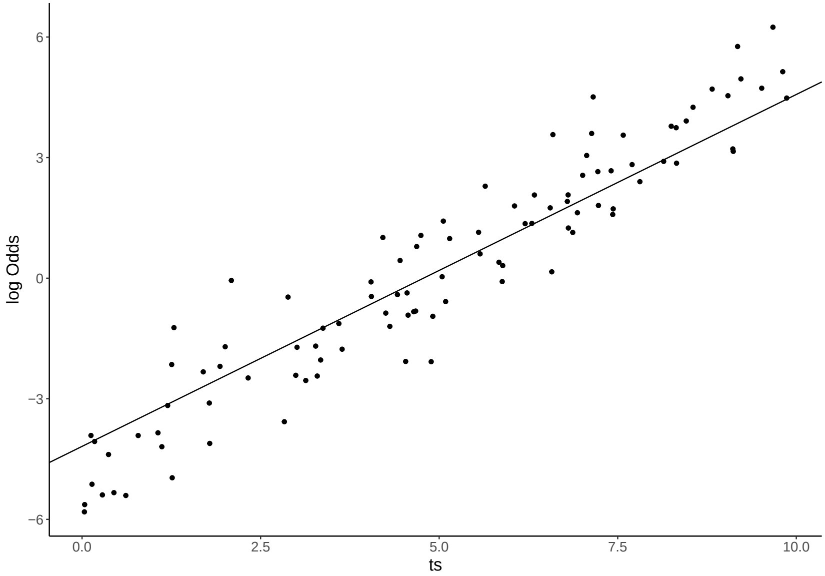
```r
set.seed(934875)
n <- 100
dat <- data_frame(ts = runif(100, 0, 10), # tumor size
                  lodds = rnorm(n, (ts - 5)), # log odds of metastisis
                  odds = exp(lodds),
                  p = odds / (1 + odds),
                  metastasis = rbinom(n, 1, p))

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