STAT 232 Categorical Data Analysis **R commands**

(1) To compute $p = P(\chi^2 > a)$ with ν degrees of freedom, use

$$1 - \text{pchisq}(a, \text{df} = \nu) \;\; \text{or} \;\; \text{pchisq}(\,a,\, \text{df} \,=\, \nu, \text{lower.tail} = \text{FALSE})$$

(2) To compute $\chi^2_{\alpha,\,\text{df}=\nu}$, use

$$\text{qchisq}(1 - \alpha, \text{df} = \nu) \;\; \text{or} \;\; \text{qchisq}(\alpha, \text{df} = \nu, \text{lower.tail} = \text{FALSE})$$

(3) Let Z be a standard normal variable. To compute, $P(Z \geq a)$, use

$$1 - \text{pnorm}(a, \text{mean} = 0, \text{sd} = 1) \;\; \text{or} \;\; \text{pnorm}(a, \text{mean} = 0, \text{sd} = 1, \text{lower.tail} = \text{FALSE})$$

(4) Let Z be a standard normal variable. To compute, $P(Z \leq a)$, use

$$\text{pnorm}(a, \text{mean} = 0, \text{sd} = 1) \;\; \text{or} \;\; \text{pnorm}(a, \text{mean} = 0, \text{sd} = 1, \text{lower.tail} = \text{TRUE})$$

(5) Let Z be a standard normal variable. To compute, $P(Z \leq a) = p$, where p is given, use

$$\text{qnorm}(p, \text{mean} = 0, \text{sd} = 1) \;\; \text{or} \;\; \text{qnorm}(p, \text{mean} = 0, \text{sd} = 1, \text{lower.tail} = \text{TRUE})$$

(6) Let Z be a standard normal variable. To compute, $P(Z \geq a) = p$, where p is given, use

$$-\Big(\text{qnorm}(p, \text{mean} = 0, \text{sd} = 1)\Big) \;\; \text{or} \;\; \text{qnorm}(p, \text{mean} = 0, \text{sd} = 1, \text{lower.tail} = \text{FALSE})$$

(7) θ has a gamma distribution with parameters α and β. To compute $P(\theta \geq a) = p$, use

$$\text{qgamma}(1 - p,\ \alpha,\ 1/\beta)$$

(8) θ has a gamma distribution with parameters α and β. To compute $P(\theta \leq a) = p$, use

$$\text{qgamma}(p,\ \alpha,\ 1/\beta)$$

(9) Let Y be a beta-distributed random variable with parameters α and β. To compute $P(Y \leq y_0)$, use

$$\text{pbeta}(y_0,\ \alpha,\ 1/\beta)$$

(10) Let Y be a beta-distributed random variable with parameters α and β. To compute the pth quartile a, that is, to find the value of a such that $P(Y \leq a) = p$, use

$$\text{qbeta}(p,\ \alpha,\ 1/\beta)$$

(11) How to find and a 95% Bayesian credible interval for the difference in proportions?

$>$ num.samp = 1000

$>$ p1 = rbeta(num.samp, enter the α value, enter the β value) # sample from science

$>$ p2 = rbeta(num.samp, enter the α value, enter the β value) # sample from humanities

$>$ quantile(p1 - p2, c(0.025, 0.975)) #95% CI for difference in proportions

(12) How to find a 95% CI for the odds ratio based on your posterior distribution.

$>$ or $<-$ $(p1/(1 - p1))/(p2/(1 - p2))$ # find odds ratio for each sample

$>$ quantile(or, $c(0.025, 0.975)$) # gives 95% credible interval for or

(13) How to find the posterior probability that one group is more likely to take a statistics course than a humanities major?

$>$ mean($p1 > p2$) # approx. posterior $P(p1 > p2)$, (science $>$ humanities)