

## Final exam

- 2.5 hours allotted
- Chapters 1-17, 19, 20
- **Excluding** section 9.3, 9.6, 14.7, 15.3, 15.6, confidence interval for  $r$
- Bring: Calculator (*not* programmable), pen or pencil, UBC ID
- You will be given: formula sheet & stats tables

## One variable: Which test?

Data type	Goal	Test
Categorical	Comparing a proportion to a hypothesized value	Binomial test (7)
		$\chi^2$ Goodness-of-fit test with two categories (used if sample size is too large for the binomial test) (8)
	Comparing frequency data to a probability distribution	$\chi^2$ Goodness-of-fit test (8)
Numerical	Comparing mean to a hypothesized value when data are approximately normal (possibly only after a transformation) (13)	One-sample $t$ -test (11)
	Comparing median to a hypothesized value when data are not normal (even after transformation)	Sign test (13)
	Comparing frequency data to a discrete probability distribution	$\chi^2$ Goodness-of-fit test (8)
	Comparing data to the normal distribution	Shapiro-Wilk test (13)

## Two variables: Which test?

		Type of explanatory variable	
		Categorical	Numerical
Type of response variable	Categorical	Contingency analysis (9)	Logistic regression (17)
	Numerical	$t$ -tests, ANOVA, Mann-Whitney $U$ -test, etc. [See Table 3 for more details.]	Linear and nonlinear regression (17)  Linear correlation (16) Spearman's rank correlation (when data are not bivariate normal) (16)

## Tests to know how to do...

- Binomial test
  - Directly
  - Using normal approximation
- $\chi^2$  Goodness-of-fit test
- $\chi^2$  Contingency analysis
- One-sample  $t$ -test
- Paired  $t$ -test
- Two-sample  $t$ -test
- Sign test
- Single factor ANOVA
- Correlation
- Linear regression

## Tests to know about...

- Fisher's exact test
- Welch's  $t$ -test
- Shapiro-Wilk test
- Mann-Whitney U test
- Tukey-Kramer test
- Kruskal-Wallis test
- Multifactor ANOVA
- Spearman's correlation
- ANCOVA
- Polynomial regression
- F-tests and Levene's test
- Logistic regression
- Randomization
- Simulation
- Bootstrapping
- Likelihood

## Confidence intervals to know how to do...

- Proportion
- Mean of a normally distributed variable
- Variance of a normally distributed variable
- Regression slope

## Know tests to...

- Compare proportion to constant
- Compare proportion from two groups
- Test independence of two categorical variables
- Compare frequency data to a model
- Compare mean to a constant
- Compare means of two groups
- Compare mean difference of two groups (paired data)
- Compare means of more than two groups
- All of the above, *without* assuming normal distribution

## Know tests to...

- Test for independence of two numerical variables - with assumptions
- Test for independence of two numerical variables - with fewer assumptions
- Predict one numerical variable from another
- Compare two slopes
- Test the interaction of two categorical factors effects on a numerical variable factor
- Compare to a normal distribution
- Compare the variances of two groups
- Predict a binary variable from a numerical variable

## Know tests to...

- Deal with new situations, not dealt with by traditional tests

## Practice - which test?

1. Do people who receive a vaccine differ from those who do not in whether they get a disease in the next 5 years?
2. Do people who finish college on average get a higher income than people who do not?

## Practice - which test?

3. How can we use the wing length of a bumble bee to predict its maximum flying speed?
4. How much variation in flying speed of bumblebees is predicted by their wing lengths?
5. Is the weight at age 2 different on average for sets of dogs that are fed one of 5 different kinds of dog foods?

## Practice - which test?

6. Does the number of yellow/wrinkled, yellow/smooth, green/wrinkled, and green/smooth peas in a cross fit the 9:3:3:1 ratio predicted by Mendel?
7. Does the number of green and yellow peas from a cross fit the 3:1 ratio predicted by Mendel? (Imagine you have only 17 data points.)

## Practice - which test?

8. Does the growth rate of sea stars vary with temperature? (with and without particular assumptions)

9. Does the relationship of sea star growth rate and temperature vary depending on whether calcium is added to the water or not?

## Practice - which test?

10. Is the mean length of elephant trunks different in males and females?

11. Are people equally likely to be born on each of the 7 days of the week?

## Practice - which test?

12. An experiment measured the effects of a treatment adding calcium to a diet, a treatment adding selenium to a diet, adding both, or neither, and measured the swimming speed of fish. Is there an effect on swimming speed of adding calcium, adding selenium, or an interaction between those two?

13. Is tree height normally distributed?

14. Do trees with and without added fertilizer have the same variance in growth rate?

## Practice - which test?

15. Can we predict whether seed germinates or not from the weight of the seed?

16. The developmental pathway leading to the formation of spots on butterfly wings has been studied by surgical excision of a small amount of tissue on the left wings of a set of butterflies, with the right wings left untouched. The size of the spots on these wings was subsequently measured. How would you test whether the manipulation had an effect on spot size?

## Practice - which test?

- 1. In the World Series, the teams have won at home 286 times and lost at home 243 times. Is there a "home-field advantage"?
- 2. Are the mean numbers of parasites per host different in multiply-infected vs. singly-infected hosts? (Answer if number of parasites per host is normally distributed, and if it is not.)

## Practice - which test?

- 3. Are longevity and fecundity related?
- 4. The number of eggs per worm seems to decline in a curved relationship with worm density. How can you test this relationship?
- 5. Longevity (number of days until death) in response to cold temperature may be a function of temperature or of genotype. How can you test whether these two effects are independent?

## Practice - which test?

- 6. The developmental pathway leading to the formation of spots on butterfly wings has been studied by surgical excision of a small amount of tissue on the left wings of a set of butterflies, with the right wings left untouched. The size of the spots on these wings was subsequently measured. How would you test whether the manipulation had an effect on spot size?
- 7. Does human height follow a normal distribution?
- 8. Is the variance of male weight equal to the variance in female weight (if weight is not perfectly normal)?

## Practice - which test?

- 9. Five different kinds of dog food were each tested on 25 young dogs, to determine which made them grow most quickly (as measured by g/day/kg). Which test(s) might you use?
- 10. Five different kinds of dog food were each tested on 25 dogs, to determine whether they had an effect on whether or not the dogs developed diabetes. Which test might you use?

## Practice - which test?

- 11. 25 dogs were grown in each of five different temperatures. Is there an effect on growth rate?
- 12. 25 dogs were each grown in five different temperatures. Data were collected on whether or not the dogs survived to an age of 5 years. Is there an effect of temperature on mortality?

## Practice - which test?

- 13. An analysis of variance has successfully shown an effect of diet on the amount of bacon a pig produces. How do you determine whether a particular diet is significantly better than the others?
- 14. You discover that a data set you have addresses a very important question, but the analysis required to address that question does not fit any known existing statistical test. What are your options?