

## Statistical models & inference:

*estimate and predict uncertain quantities, events*

### Probability: Measurement of uncertainty

Interpretations?

- **Counting**: equally likely outcomes (card and dice games, sampling)
  - **Frequency**: rain tomorrow? Relative frequency of rain on “similar” days
  - **Subjective**: measure of belief, judgement
    - includes counting, frequency cases, BUT:
    - unique events – Duke:UNC BBall on Jan 27th
    - **conditioning**: rain tomorrow conditional on – ?
- Rain today ? WeatherChannel forecast? When? Who?

# PROBABILITY AND INFERENCE: INTRODUCTION

“Simplest” statistical inference context: DIAGNOSIS example

- $x = 0, 1$ 
  - Data, to be observed, result of experiment/observation
  - (Binary) Sample Space  $x \in \{0, 1\}$  – discrete, simple
- $\theta = 0, 1$ 
  - Parameter, “Truth,” state of nature, uncertain/unknown
  - Primary objective: Inference on  $\theta$
  - (Binary) Parameter Space  $\theta \in \{0, 1\}$  – discrete, simple
- Uncertainty about outcomes described by **conditional probabilities**  $p(x|\theta)$

NOTATION:  $p(x|\theta)$ ,  $p(x = 1|\theta = 1)$  etc,  $Pr(x = 1|\theta = 1)$

- $p(x|\theta)$  defines the **Sampling model**

Where do these probabilities come from?

Assumptions about **Population** and **Random Sampling**

- Measures of **likelihood** of observed data given states of nature
  - $x = 1$  is 9.9 times as likely when  $\theta = 1$  than when  $\theta = 0$
  - Likelihood ratios  $r(x) = p(x|\theta = 1)/p(x|\theta = 0)$
  - $r(1) = 9.9$ ,  $r(0) = 0.01111\dots$
- Evidence, Information to compare  $\theta = 1 : 0$  given data  $x$

What is the probability of disease?

- Which probability?
- Are “you” diseased?
- Based on what evidence and information?

### Initial probability

Based on all past data, assumptions  $\rightarrow$  “information set”  $H$

- Take this as the population disease incidence rate ( $/100$ )
- Initial probability = **Prior probability**  $p(\theta = 1)$
- Formally, explicitly:  $p(\theta = 1) \equiv p(\theta = 1|H)$ 
  - Where does this probability come from?
  - Assumptions about **Population** and **Random Sampling**

- Odds on disease (initial or prior odds), **Odds = Prob/(1-Prob)**

or

$$o(\theta = 1) = \frac{p(\theta = 1)}{(1 - p(\theta = 1))} = \frac{p(\theta = 1)}{p(\theta = 0)}$$

- **Prob = Odd/(1+Odds)** or  $p(\theta = 1) = o(\theta = 1)/(1 + o(\theta = 1))$
- e.g.:  $p(\theta = 1) = 0.01$  and  $o(\theta = 1) = 0.0101\dots$

## JOINT PROBABILITY DISTRIBUTION: $p(x, \theta) = p(x|\theta)p(\theta)$

- Marginal distribution for  $x$  via *Law of Total Probability*:

$$p(x) = p(x|\theta = 0)p(\theta = 0) + p(x|\theta = 1)p(\theta = 1)$$

- **Bayes' theorem:**

$$p(\theta|x) = \frac{p(\theta)p(x|\theta)}{p(x)}$$

- Bayes' theorem reverses conditioning
- Now  $x$  is fixed, conditioned upon
- Initial/prior probability distribution for  $\theta$  is revised/updated
- Final probability = **Posterior probability**  $p(\theta|x)$

In odds form,

$$o(\theta = 1|x) = o(\theta = 1)r(x)$$

based on observed likelihood ratio  $r(x) = p(x|\theta = 1)/p(x|\theta = 0)$

Posterior odds = Prior odds  $\times$  Likelihood ratio

- Scientific learning/Statistical Inference:  $p(\theta) \rightarrow p(\theta|x)$
- Mapping via **likelihood function**  $p(x|\theta)$  from sampling model
- Formally,  $p(\theta|H) \rightarrow p(\theta|H, x)$  where  $H$  = prior information
- *Technically*: Simply apply Bayes' theorem

“Final” inferences depend on both data  $x$  and prior information  $H$

**Sensitivity to prior:**

- Any prior  $a = p(\theta = 1)$  implies posterior  $p(\theta|x)$
- Variation with respect to  $a$  for given data, hence fixed  $r(x)$  ... ?
- graph  $p(\theta = 1|x)$  versus  $a$

