Prediction

Todd Davies

Symsys 130

April 22, 2013

Variable Prediction

Independent variable (IV) – predictor

Dependent variable (DV) - target

Judgment and Prediction

Estimation – predict a population statistic from a sample statistic for the same variable

- Sample mean -> Population mean
- Sample frequency -> Long-term frequency
- Subjective probability -> Actual probability
- Intuitive correlation -> Pearson correlation

Cross-variable prediction – predict a value of one variable based on another

- Prior odds, likelihood ratio -> Posterior odds
- Predictor variables -> Target variable in regression model

Perspectives on human judgment

Heuristics and Biases (HB, Tversky, Kahneman)

Other approaches

- Rational analysis of cognition (Chater, Oaksford)
- Bounded rationality (Gigerenzer, Hertwig)
- Naturalistic Decision Making (NDM, Klein, Shanteau)

Intuitive correlation

Jennings, Amabile, and Ross (1982) gave subjects lists of number pairs. The task was to assign a number between -100 and 100 (inclusive) to each list based on the strengt hof the relationship between the pairs. Quoting Baron (2000): “The true correlation coefficients of the numbers in each list differed from list to list: the range was from 0 to 1. Subjects gave ratings near 100 for correlations of 1, and ratings near 0 for correlations of 0. For correlations of 5, subjects gave ratings of about 20. These results tell us that the naïve idea of 'degree of relationship', although it resembles correlation, is not quite the same as the relationship measured by the correlation Coefficient. The subjects' deviation from mathematical correlation was not an error, for the correlation coefficient is only one of many possible measures of association, and there is no reason subjects should use this particular measure.”

Illusory correlation

Chapman and Chapman (1971) gave college students drawings from the Draw-a-Person test, a tool used in clinical diagnosis of mental disorders. Each drawing was labeled with a psychological characteristic supposedly corresponding to the person who made the drawing. Example characteristics were “suspicious of other people” and “has had problems of sexual impotence”. The labels were chosen so that there was no correlation between psychological characteristics and pictorial features in the drawings widely believed to be associated with these characteristics, e.g. “big eyes” for “suspicious of other people” and sexual features for “has had problems of sexual impotence”. The subjects were asked to discover relationships between drawing features and psychological characteristics from these drawings, and reported the correlations that widely held prior beliefs predicted would exist. The authors called this “illusory correlation”. The correlations that were found by subjects in the Draw-a-Person test experiment were the same as ones believed by clinicians who worked with actual patients, but no such correlations existed in the patient population. A similar study was done with Rorschach inkblot test data. Students found illusory positive correlations in Rorschach data between patients' actual responses and their clinical diagnoses when a relationship between the response and diagnosis is naively predicted. They also *failed* to find relationships that did exist when these were *not* expected.

The hot hand – flipside of the gambler's fallacy

Gilovich, Vallone, and Tversky (1985, handed out) recruited 100 basketball fans from the student bodies of Stanford and Cornell Universities to fill out a questionnaire. The sample included 50 captains of intramural basketball teams, and all subjects played basketball at least "occasionally" (65% played "regularly"). All watched at least 5 games per year and 73% watched more than 15 games a year. The authors write:

The fans were asked to consider a hypothetical player who shoots 50% from the field. Their average estimate of his field goal percentage was 61% 'after having just made a shot,' and 42% 'after having just missed a shot.' Moreover, the former estimate was greater than or equal to the latter for every respondent. When asked to consider a hypothetical player who shoots 70% from the free-throw line, the average estimate of his free-throw percentage was 74% "for second free throws after having made the first," and 66% "for second free throws after having missed the first." Thus, our survey revealed that basketball fans believe in "streak shooting."

Gilovich et al. collected data from National Basketball Association (NBA) games for both field goals and free throws. In both cases, they found no evidence for streak shooting, also known as the "hot hand" hypothesis. Serial correlations (the correlations between boolean outcomes on a shot given the previous shot) were negative on average, the opposite of what people predict, and there were no more streaks in the shooting data than would be expected by chance.

Confusing likelihood for posterior odds

What does it mean to say that marijuana is a gateway drug?

Calibration

Probabilistic confidence = |Assessed probability – 0.5|

Confidence intervals

Common findings about overconfidence