

Decision tools

◇ Chapter 21: Decision analysis

When you have to make a choice and don't make it,
that is in itself a choice.

William James

dr. s.kaufman, levin college, csu

Decision tools

- ◇ Review, Chapter 13
- ◇ Chapter 21: Decision analysis

dr. s.kaufman, levin college, csu

Association statistics

- ◊ Chapter 13:
 - Chi Square nonparametric test
 - ◊ Independence tests
 - ◊ Goodness of fit tests

dr. s. kaufman, levin college, csu

Association – qualitative data

- Qualitative data:
 - examine observed frequencies & infer about frequency in the pop. of certain categories, to
 - test hypotheses about whether a model fits the data (goodness of fit test) or
 - test relationship among variable (test of independence)
 - focus on frequencies rather than values (data are nominal)
- Chi square statistic:
 - developed by Karl Pearson in 1900.
 - name derives from statistic used for tests
 - non-parametric (distribution-free)

dr. s. kaufman, levin college, csu

Goodness of fit

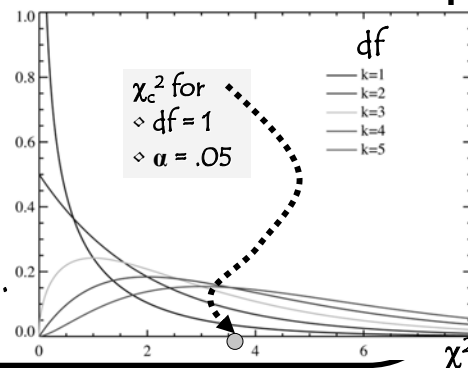
- How well does a statistical model fit a sample of *count* data? We need:
 - an appropriate statistical model
 $DATA = FIT + RESIDUAL$
 - the ability to fit the model (adjustable parameters)
 - the ability to assess the model's goodness of fit
- The FIT portion, or model
 - has adjustable parameters, so the FIT can come close to the DATA.
 - we attempt to make the RESIDUALS as small as possible within the model. →

Between two models,
the one with smaller RESIDUALS fits the DATA better.

dr. s. kaufman, levin college, csu

The chi square distribution

- Notation: χ_{df}^2 (df= degrees of freedom)
- There is one curve for each number of df
 - Die example: $df = 6-1 = 5$
 - The higher the # of df, the flatter the curve, and the more centered to the right
 - The area underneath is 1.



dr. s. kaufman, levin college, csu

Chi square distribution properties

- χ^2 is never negative (can be zero)
- Distribution: asymmetrical
 - skewed to the right
- Expected frequencies in a cell should be at least 5
- Test of *independence* – $df=(r-q)(c-1)$
 - determine if pattern of frequencies is the same for each variable as overall \rightarrow then variables are independent
- The *goodness-of-fit* test: $df = k-1$ for k categories
 - determine whether observed frequencies of a phenomenon cast doubt on a model of it:

DATA = FIT + RESIDUALS
the test tells if residuals are so large
that we should reject the model.

dr. s.kaufman, levin college, csu

Chi square test nuts and bolts

- First step: select the significance level
- Use *frequencies*, not *percentages*
- Expected frequencies need not be integers
- Cells must be *mutually exclusive* and *exhaustive* with respect to values of the variables studied
- The χ^2 distribution has an additive property \rightarrow
 - χ^2 can be pooled, or combined into one value based on χ^2 results from independent experiments;
 - the new χ^2 – the sum of all separate χ^2 – has an approximate χ^2 distribution with
degrees of freedom = sum of all df from each independent experiments.

dr. s.kaufman, levin college, csu

Chi square test, hypotheses

- Test of goodness of fit:
 - H_0 : the data fits the model
 - H_1 : the data does not fit the model
(residuals too large)
- Test of independence:
 - H_0 : the variables are statistically independent
(one does not hold information about the other)
 - H_1 : the variables are not independent

dr. s. kaufman, levin college, csu

Uses of the independence test

- Assessing test validity:
 - consider a test supposed to discriminate between two states, e.g.:
 - - on/off
 - diagnostic/actual state
 - If the two data streams are independent → the test is not valid
- Assessing test reliability:
 - compare outcomes of applying same test to the same population twice
 - If the two data streams are independent → the test is not reliable

dr. s. kaufman, levin college, csu

Example notes - proportions

- A comparison of proportions is a *goodness of fit test*, with the null hypothesis that *all proportions are equal*

- The *equal proportions model* has to be translated into frequencies for the test:

$$p \times n_i$$

- The degrees of freedom for $2r$ comparisons, where r = the number of observed values O_i and also the number of expected values E_i :

$$df = r - 1$$

dr. s.kaufman, levin college, csu

r x c contingency tables for independence tests

- Test the independence of row from column variable; labels matter:

χ^2 computation has a direction

- Identify the row variable as criterion (r categories), the column variable as predictor (c categories)
- Use row marginals to proportions to compute expectations
- If data are percentages or proportions, convert to numbers
- H_0 : row variable frequencies not affected by column variable (independence)

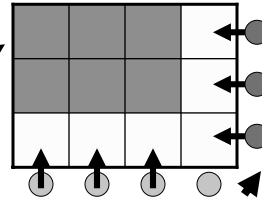
No causality!!!

dr. s.kaufman, levin college, csu

r x c contingency tables for independence tests

- Test using χ^2 with $df = (r-1)(c-1) \rightarrow$ for df number of cells compute:

$$\text{Expected (cell}_{ij}) = \text{total} \times (\text{total row}_i / \text{total}) \times (\text{total column}_j / \text{total})$$



- For example, for a 3 x 4 table, compute only $(3-1)(4-1) = 6$ cells
- The rest can be obtained as differences between computed cells and marginals

dr. s. kaufman, levin college, csu

SPSS example

- Is the verbal IQ independent of the performance IQ?

Descriptive Statistics					
	N	Mean	Std. Deviation	Minimum	Maximum
Verbal IQ Score	40	112.35	23.616	71	150
Performance IQ Score	40	111.03	22.471	72	150

Chi-Square Test

Why?

Test Statistics		
	Verbal IQ Score	Performance IQ Score
Chi-Square ^{a,b}	18.000	19.850
df	19	20
Asymp. Sig.	.522	.467

What do you conclude?

What does that mean?

- a. 20 cells (100.0%) have expected frequencies less than 5. The minimum expected cell frequency is 2.0.
- b. 21 cells (100.0%) have expected frequencies less than 5. The minimum expected cell frequency is 1.9.

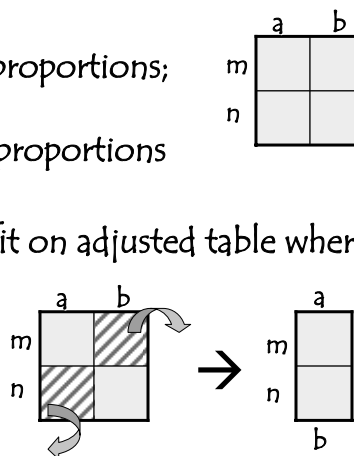
Test of non-independent proportions

- If proportions are based on samples that are
 - Independent \rightarrow t test: $H_0: p_1 = p_2$
 - Non-independent \rightarrow χ^2 test: $H_0: p_1 = p_2$ (goodness of fit)
 - We cannot distinguish a difference in *levels* (proportions of the population), since there is *overlap* of groups
 - To determine if there is a *difference in level* (not to determine the level!!!), *the overlap group must be eliminated* and the test rerun with individuals falling in only one r and one c category.

dr. s. kaufman, levin college, csu

Compare proportions

- 2 steps:
 1. Test independence of the proportions; if independent \rightarrow use t-test of difference of proportions
 2. If not independent, use t-test of goodness of fit on adjusted table where overlap is eliminated:



dr. s. kaufman, levin college, csu

Variance sampling distribution

- The sampling distribution of the variance of a normally distributed variable:

$$(n-1) \frac{s^2}{\sigma^2} \sim \chi_{n-1}^2$$

dr. s. kaufman, levin college, csu

Association statistics

- ◇ Chapter 21:
Decision analysis

- ◇

dr. s. kaufman, levin college, csu

Decision making

- Why do we need to learn to make decisions?
 - Our environment is complex in several ways:
 - the issues under debate
 - scale
 - implications in time
 - interconnections between systems
 - Our capability to think through certain aspects of decision making unaided is limited:
 - cognitive shortcomings → breaks in rationality
 - capacity to handle large amounts of complex data:
 - we do not remember well and in detail
 - we are subject to predictable biases in our thinking patterns;
 - we use perceptions to come up with judgments, but our perception mechanisms are faulty and subject to bias;

dr. s. kaufman, levin college, csu

Decision making

- Resources are limited and their division is fraught with conflict:
 - we cannot agree with each others when vying for the same resources
 - our personal goals conflict with those of others and with organizational goals
- The ways we interact with people:
 - the information age--technology, transparency, volume of data, knowledge
 - the pluralistic society
 - poor interaction skills and mutual tolerance of differences
 - we interact through representatives rather than directly → we need devices to render decisions and their basis transparent and justifiable
- Inadequate information: When we
 - know the states of nature, certainty exists.
 - don't know the states of nature but we can assign probabilities to them, there is risk.
 - don't know probabilities of the states of nature, there is *uncertainty*.

dr. s. kaufman, levin college, csu

Decision making

- Under certainty assumes perfect information
 - Examples:
 - optimization of response from fire engines located at 2 sites in the city, when all problem parameters are known
 - The city of Reginald, Ohio must choose between one of 3 types of employment programs. The effectiveness of each program depends on the level of unemployment (state of nature). What are the city's options?

dr. s.kaufman, levin college, csu

Example (cont.)

- Do nothing. Ignore problems and hope the marketplace will correct them.
 - Works best if unemployment is low - the city will not spend any funds.
 - At moderate levels (5-9%) moderate costs to Reginald -- countercyclical aid and political unrest.
 - At high levels (over 9%) option 1 is disastrous, with reduced demand for goods catapulting Reginald into a major recession.
- Operate a job placement center to find jobs for those currently out of work.
 - At low levels, minor fixed costs.
 - At moderate levels, major benefits, since sufficient jobs are available to place most temporarily out-of-work people.
 - At high levels, program is inadequate.
- Implement a job training program.
 - Below 5% unemployment, high fixed costs make it unattractive.
 - At 5-9%, minor costs.
 - Best suited for high unemployment, because it absorbs surplus labor and alters the skills of the workforce.
- Is there a scientific way to make decisions?
 - Reginald's problem can be analyzed as a decision under risk, with specific steps.

dr. s.kaufman, levin college, csu

Decision analysis

- way to decompose decisions through processes that provide
 - rationality safeguards
 - transparency for others to reconsider the problem.
- not a magic trick for identifying the best decision.
 - no guarantee that it will work, but
 - decreases the likelihood of bias and inadvertent mistakes stemming from limitations built into our thinking capability.
 - allows us to deal explicitly with issues of value, risk, uncertainty & the set of possible solutions to problems.

dr. s. kaufman, levin college, csu

Decision analysis

- Is DA best suited for some kinds of problems only?
 - not ideal for deciding value questions (abortion, religion issues, modern art or music, scientific questions).
 - helps where
 - preferences, choices and their consequences are identifiable and
 - probabilities can be attached to the consequences of choices.
- DA is an individual tool for sorting out decisions;
 - it can be turned fairly easily into a group tool (where the group has a common goal and debates the best way to reach it).
 - It is ill suited for devising a settlement in a dispute, although each group or individual involved can use DA to generate choices in the dispute.

dr. s. kaufman, levin college, csu

Approach

- DA decomposes any decision situation into:
 - preferences--how would an individual order the possible outcomes of the decision? (depends on values)
 - choice set--the set of actions seen as available to the decision maker (e.g., several makes of cars, several sizes of budgetary allocations to a particular program, etc.)
 - contingencies--states of nature and their likelihoods (e.g., a downturn in the economy, a successful technology, etc.)
 - probabilities for consequences of choices--the likelihoods that a set of consequences will follow the decision (e.g., frequency of repairs by car model, probability that certain amounts of spending will accomplish goals, etc.)
 - decision rule

dr. s.kaufman, levin college, csu

Decision rule

- based on the *expected* benefits of the ultimate choice.
 - combine the values of various outcomes with their probabilities of materializing → select the outcome with:
 - the largest expected value, considered the most rational choice, maximizing expected utility, with advantages and disadvantages.
 - the smallest maximum loss (*minimax*)
 - the largest minimum gain (*maximin*)
 - The smallest the minimum loss (*minimin*)
 - → can go back and work with the decision inputs and explore the consequences of changing preferences or the likelihoods of consequences (*sensitivity analysis*);
 - → as representative of a constituency, can explain what your decision was based on, or customize it to the preferences of the constituency;
 - Problems:
 - ignores all consequences that are possible but not known
 - difficulty incorporating non-quantifiable items and items for which the probability distribution is unknown.

dr. s.kaufman, levin college, csu

Example

- The city of Reginald, Ohio must choose between one of 3 types of employment programs. Which to choose?
- DA steps:
 1. Set up a decision table

Decision option	Unemployment level		
	Low	Moderate	High
Do nothing			
Placement			
Job training			

dr. s. kaufman, levin college, csu

Example (cont.)

2. Calculate the payoff to the city for each cell in the table, using past experience, cost-benefit calculations, etc.
 - Joyce Carruthers, mayor of Reginald, has asked Thomas Malthus, a labor economist at Reginald State U., to calculate costs and benefits.

Decision option	Unemployment level (K = \$1,000)		
	Low (.2)	Moderate (.3)	High (.5)
Do nothing	\$20K	-\$5k	-\$120K
Placement	-\$10K	\$50K	-\$30K
Job training	-\$110K	\$10K	\$140K

dr. s. kaufman, levin college, csu

Example (cont.)

3. Determine the probability that each state of nature (here, unemployment level) will occur.

- For the period covered by the decision, what is the probability of low, moderate or high unemployment? Calvin Kent, local economics consultant, estimated the probabilities

4. Calculate the expected value of each option:

$$EV = \sum p_i v_i$$

$$20 \times .2 + (-5) \times .3 + (-120) \times .5 = -57.5K$$

dr. s.kaufman, levin college, csu

Example (cont.)

Decision option	Unemployment level			EV
	Low (.2)	Moderate (.3)	High (.5)	
Do nothing	\$20K	-\$5k	-\$120K	-\$57.5K
Placement	-\$10K	\$50K	-\$30K	-\$2K
Job training	-\$110K	\$10K	\$140K	\$51K

5. Malthus recommends the strategy that carries the highest expected value (which is it here?)
What should Joyce Carruthers do?

dr. s.kaufman, levin college, csu

Decisions under risk – Assumptions

- The decision maker has the same attitude toward risk regardless of the size of the benefits involved
 - risk neutrality, as opposed to risk aversion or risk seeking:
The expected value of a decision with .25 probability of a \$10,000 loss is the same as that of a decision with a probability of .0025 for a \$1 million loss;
this assumption says we could choose either.
- The expected value covers all the values the decision maker wants to maximize.
 - One may want to incorporate different values and political criteria.

dr. s.kaufman, levin college, csu

Example (cont.)

- Mayor Caruthers is visited by J. Barringford Tipton of Chaste Econometrics, a well-known forecasting firm.
 - Tipton tells Caruthers that Chaste can accurately forecast the Reginald's unemployment rate the next year, for \$25,000.
 - Should Caruthers hire Chaste?
- What is the value of perfect information?
(knowing the state of nature in advance)
 - Without perfect information, Caruthers would select option III, with an EV of \$51,000.
 - Compare this figure with the expected value if the city had perfect information.

dr. s.kaufman, levin college, csu

Example (cont.) – value of information

- If the state of nature was
 - *low* (20% of the time) → Reginald would do nothing and would reap a benefit of \$20K;
→ the expected value of this occurrence is $20K \times .2$;
 - *moderate* (30% of the time) → Reginald could do placement and would get a benefit of \$50K;
→ the expected value of this occurrence is $50K \times .3$;
 - *high* (50% of the time) → Reginald could do job training to get a benefit of \$140K;
→ the expected value of this occurrence is $140K \times .5$;
- $EV = .2 \times 20K + .3 \times 50K + .5 \times 140K = \$89K$
 $EV_{\text{perfect info}} - EV_{\text{no info}} = \$89K - \$51K = \$38K > 25K$

dr. s. kaufman, levin college, csu Should Carruthers buy?

Decision analysis (DA) uses

- Situations with numerous nodes and intricate probability distributions:
 - different measurement scales (not \$)
 - lack of knowledge about some probabilities
 - high stakes
 - information is complex
- then DA is an ideal tool
 - usefulness decreases as we get less precise in
 - specifying preferences
 - giving probabilities,
 - identifying all the members of our action space.
- Other techniques -- brainstorming, Delphi, etc.

dr. s. kaufman, levin college, csu

Decision trees

- Under complex conditions, represent decisions with *decision trees*.
 - Example: deployment of the MX missile
 - decision nodes
 - chance nodes
- Resolution of decision trees: backwards from branches to trunk
 - decision rule is the same as for contingency table (select strategy with highest expected value)
 - for highly complex trees, only computer solutions.

dr. s. kaufman, levin college, csu

Decision tree components

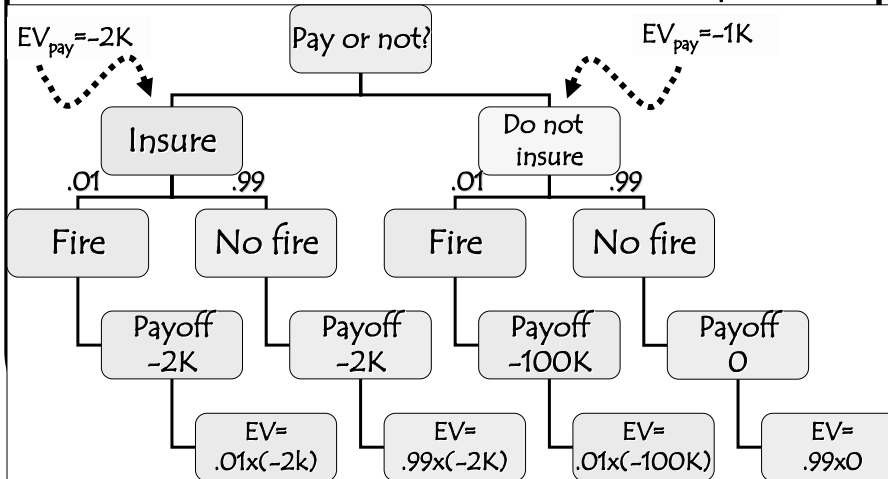
- Decision alternatives a_i , possible actions
- Outcomes:
 - the consequences of each action, each with a certain probability of materializing
→ same as the sample space for a probability experiment
- Outcome probabilities, p_i , depending on
 - our choices or
 - states of nature over which we have no control (fire or earthquake compared to investment yields)
- Outcome payoff values x_i ,
 - the values to us of outcomes materializing; they have to be ratio-scaled for DA
- Expected value of a choice of action a_i is:

$$E(a_i) = \sum_{(1 \text{ to } n)} p_i x_i$$

dr. s. kaufman, levin college, csu

Example:

- Would you pay \$2K for a fire insurance policy on a remote and vacant warehouse valued at \$100K with a fire prob. = .01?



Decision making under uncertainty

- Probabilities for states of the world are not known. What can be done?
 - Bayesian approach: incorporate subjective judgments (knowledge and intuition)
 - Insufficient reason approach: assume events are equally probable in the absence of better information (no reason to believe otherwise)
 - Maximin principle: select best of worst cases
 - Minimax regret: minimize opportunity costs of wrong decision
 - Maximax: assume the best will happen
- Select decision rule based on:
 - how important gains and losses are,
 - what the consequences would be to how many people, etc.

dr. s.kaufman, levin college, csu

Game theory

- In some situations the states of nature are replaced by the decisions made by other *actors* → game
- Game theory analyzes the strategies to select in situations where the outcome depends on what others decide.
 - zero-sum: one player's gain is the other's loss
 - positive-sum games: opportunities for win-win
 - the prisoners' dilemma → the tragedy of the commons

dr. s. kaufman, levin college, csu