Bienvenue

Willkommen

Welcome

welkom

yôkoso

Benvenuto

Bienvenida

tervetuloa

รับเสด็จ

欢迎

LECTURE NO. 3
Statistical Testing

SVI Endowed Chair for
International Direct Marketing
DMCC – Dialog Marketing Competence Center
University of Kassel
rwagner@wirtschaft.uni-kassel.de

February 7th , 2014

DMCC DIALOG MARKETING COMPETENCE CENTER

AGENDA

1. Introduction
2. Research Methods
3. Theory Building & Hypotheses
4. Selecting Statistical Tests
5. Common Statistical Errors
6. Excursus: Nothing can be Proved Empirically

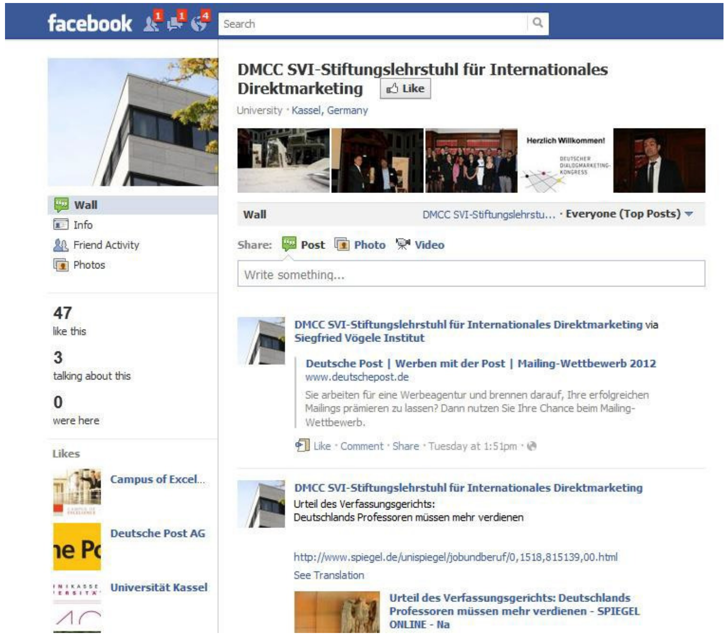
1. INTRODUCTION

SVI Endowed Chair for International Direct Marketing

Team

Team: Carina Erk, Taylan Urkmez, Jasmin Dittmar, Kristina Richter, Katrin Zulauf, Dr. Christine Falkenreck, Christina Sänger, Nicolai Mollee, and Prof. Dr. Ralf Wagner

SVI Endowed Chair for International Direct Marketing



Latest News

Twitter:	http://twitter.com/DMCC_Wagner
Website:	www.dmcc.uni-kassel.de/wagner
Facebook:	DMCC-SVI

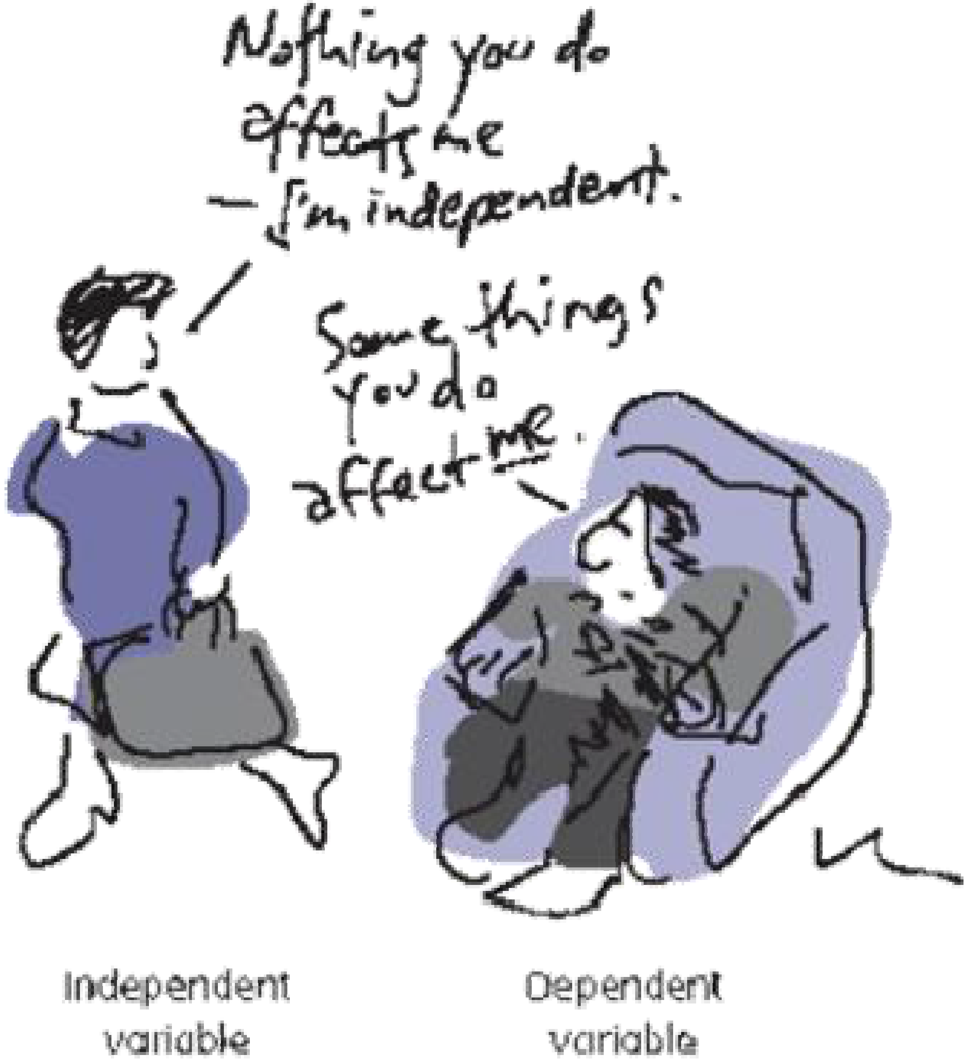
Curriculum Vitae Prof. Dr. Ralf Wagner

- Since 2006 professor at the University of Kassel, Germany
- Holding the SVI endowed chair for international direct marketing
- 2000-2006 Assistant professor at Bielefeld University at the marketing department
- 1995-2000 PhD candidate at Bielefeld University at the marketing department
- Serving as visiting Professor at:
    - Academy of National Economy, Moscow
    - International Business School, University of Vilnius, Lithuania
    - University of Information Technology and Management in Rzeszow, Poland
    - Nankai University, China

2.

Research Methods

Research informs Practice, and Practice informs Research



MEANING OF RESEARCH – IN GENERAL

A scientific and systematic search for relevant information on a specific topic.

(LIMAT, n.d.)

"A systematized effort to gain new knowledge."

(Redman & Mory, 1933, p.10)

"Marketing research is the function that links the consumer, customer, and public to the marketer through information - information used to identify and define marketing opportunities and problems; generate, refine, and evaluate marketing actions; monitor marketing performance; and improve understanding of marketing as a process. Marketing research specifies the information required to address these issues, designs the method for collecting information, manages and implements the data collection process, analyzes the results, and communicates the findings and their implications."

(AMA, 2004)

Research as an Academic Activity

“Research comprises defining and redefining problems,
formulating hypothesis or suggested solutions;
collecting, organizing and evaluating data;
making deductions and reaching conclusions;
and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis.”
(LIMAT, n.d.)

RESEARCH METHODS

- Methods help to gain scientific knowledge and translate this knowledge into practical actions.
- Methods are tools for the scientific progress.
- Research methods help to solve the research problem systematically.

BUT

Source: LIMAT, n.d.; Eid et al., 2010, p. 6.

Research Methods (cont.)

BUT:

- Researchers do **not only** need to know **how** to develop certain indices/tests, calculate the mean etc., apply particular research techniques.
- They do need to know **which** of these methods or techniques, are relevant/non-relevant, and **what** would they mean/indicate.
- It is important to understand the **assumptions** underlying various techniques and to know the **criteria** by which one can decide that certain techniques/procedures will be applicable to certain problems and others will not.

Source: LIMAT, n.d.

3. Theory Building and Hypotheses

3.1 Theory Building

THEORY BUILDING

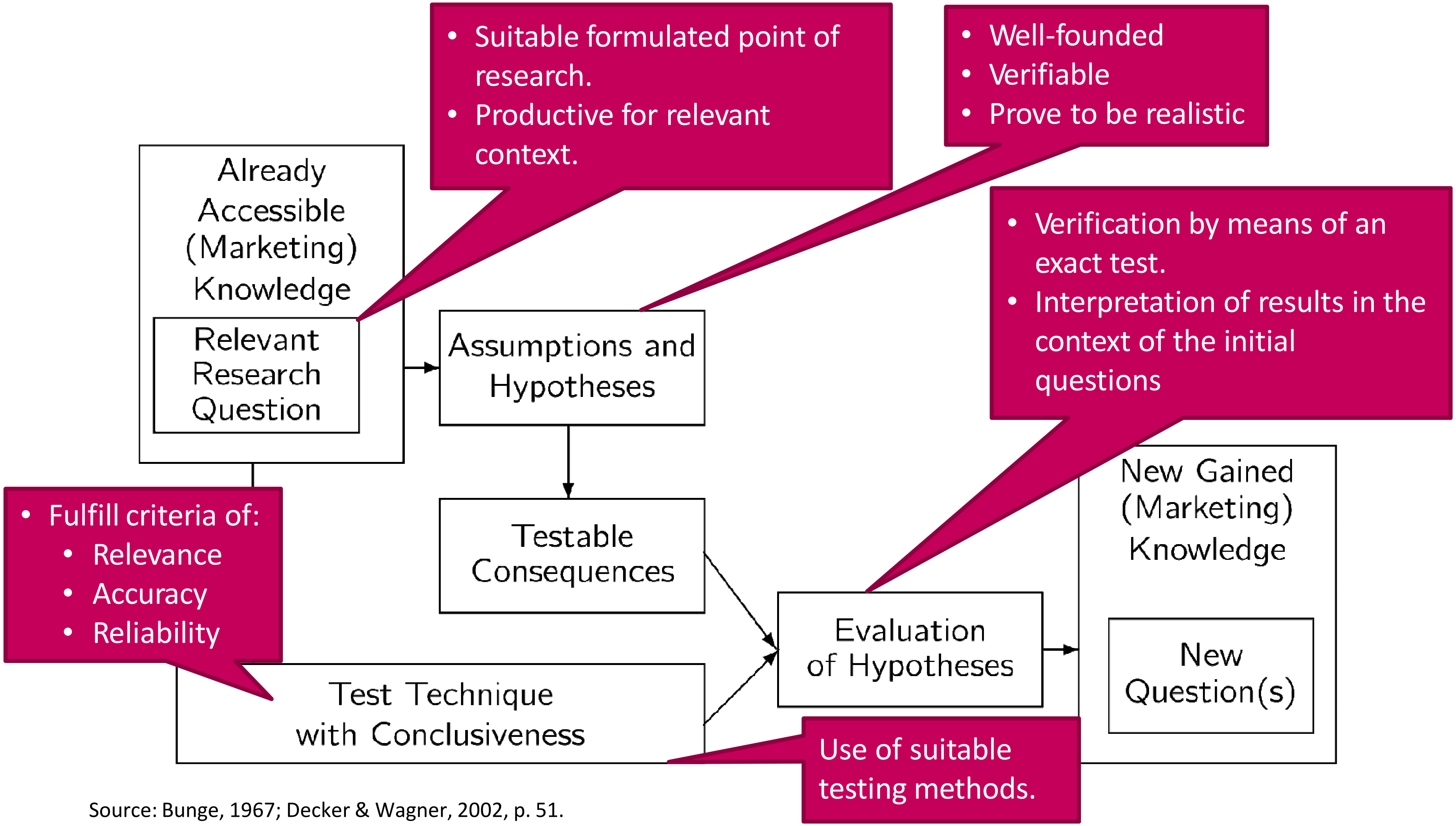
Theory Building

Theories can not be built from observable facts. Rather, they are a combination of bold statements, consistent argumentation as well as evaluation and proof by experience.
(Popper, 1989)

Significance of Theoretical Observation

Marketing research must be grounded in theory. Theory enables us to meaningfully interpret and integrate the findings with previous research. Due to an under-utilization of existing theory, our understanding of several substantive areas is limited despite numerous studies.
(Malhotra et al., 1999, p. 177)

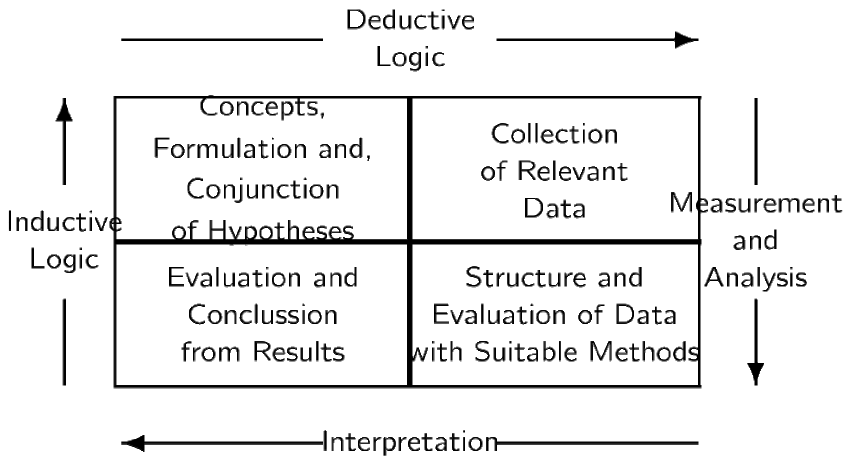
TYPICAL RESEARCH PROCESS



Source: Bunge, 1967; Decker & Wagner, 2002, p. 51.

KNOWLEDGE GAINING

Deduction: Derive a specific statement from an general one.
Induction: From a specific statement to a general one.



Double Inclusion: Valid for element *i* and element *i+1.*

Differentiation (Bass & Wind, 1995)

- Theory – Empiricism - Theory – Empiricism (TETE) (e.g., negative binomial model)
- Empiricism - Theory – Empiricism – Theory (ETET) (e.g., multinomial logit model)

Source: Decker & Wagner, 2002, p. 52.

Excursus: Binomial Model

Models of Purchase and Purchase Intention – Definitions

Behavioral Intention (AMA, 2013a)

A cognitive plan to perform a behavior or action ("I intend to go shopping later"), created through a choice/decision process that focuses on beliefs about the consequences of the action.

Purchase Intention (AMA, 2013b)

A decision plan to buy a particular product or brand created through a choice/decision process.

Types of Purchase Intentions

- Towards a complete product class
- Specific brands
- Specific brands in a specific shopping place

Excursus: Binomial Model

Models of Purchase and Purchase Intention – Definitions (cont.)

Planned Buying versus Unplanned Buying (AMA, 2013c)

Planned buying is purchasing activity undertaken with a problem previously recognized and a buying intention previously formed. Most organizational buying is planned buying. Unplanned buying is buying activity that occurs as a result of exposure to an advertisement or a salesperson's visit. Trial of new supplies or consumable supplies may begin as unplanned buying.

Question

How to come from purchase intention to purchase?

Purchase Intention

Factors of Purchase Intention

- Factors determined by individual purchase experience
- Situational factors
- Anticipative factors

Pros and Cons of Purchase Intention

- Relatively easy to examine
- Evaluation before purchase
- What about the reliability of purchase intention data?

  → Problem: randomness of buying behavior

MODEL-BASED ANALYSIS OF PURCHASE INTENTION DATA

Assumption (Kalwani & Silk, 1982)

Subjects' statements concerning the intensity of their purchase intention can be seen as the result of a random process.

K	Sample size
c	Purchase intention ($c = 0, 1, \cdots, C$)
C	Maximum category of purchase intention

0	There is no intention to buy.
1	Buying is rather unlikely.
2	They are indifferent.
3	Buying is rather probable.
4	There is a definite purchase intention.

Source: Decker & Wagner, 2002.

Calculation of Purchase Probability

1. Identification of an estimator $\hat{q}$ for “real” purchase intention q in purchase intention category c
2. Transformation from estimated “real” intention into so called unregulated purchase intention $\hat{p}_u$
3. Transformation of unregulated $\hat{p}_u$ into estimated purchase probability $\hat{p}$

Initial Point

$$\hat{q} = a + bc \qquad \text{with} \quad a, b > 0; \quad a + b < 1$$

Probability of Consumer *k* to Choose Purchase Intention Category *c*

$$P(Y_k = c) = \binom{C}{c} q_k^c (1 - q_k)^{C-c} \qquad \forall\ k$$

with q_k as (0; 1) normed "‘real’" purchase intention

DMCC DIALOG MARKETING COMPETENCE CENTER

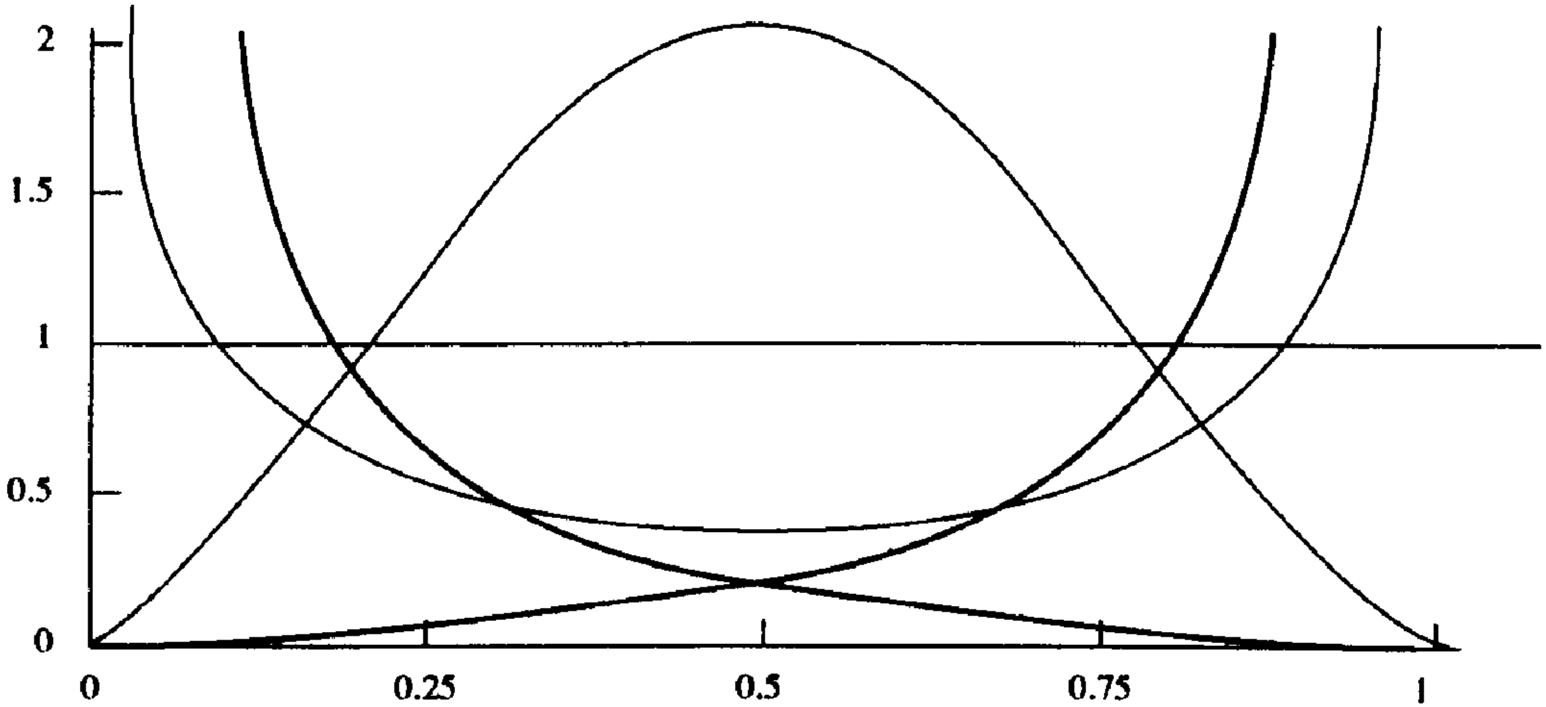
Representation of Consumer Heterogeneity

by Beta Distribution

$$f(q_k|\beta_1,\beta_2) = \frac{\Gamma(\beta_1+\beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)} q_k^{\beta_1-1}(1-q_k)^{\beta_2-1} \qquad \forall\ k$$

with $\beta_1, \beta_2 > 0$ and $\Gamma(z) = \int_0^\infty w^{z-1}\exp(-w)\mathrm{d}w$, $z > 0$;
$\Gamma(z+1) = z\Gamma(z), \quad \Gamma(1) := 1$

POSSIBLE SHAPES OF BETA DISTRIBUTIONS



Source: Lilien et al., 1992.

Beta Binomial Model for Heterogeneous Populations

$$P(Y = c) = \int_0^1 P(Y_k = c) \cdot f(q_k | \beta_1, \beta_1) \mathrm{d}q_k$$

$$= \int_0^1 \binom{C}{c} q_k^c (1 - q_k)^{C-c} \cdot \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)} q_k^{\beta_1 - 1} (1 - q_k)^{\beta_2 - 1} \mathrm{d}q_k$$

$$= \int_0^1 \underbrace{\binom{C}{c} \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)}}_{\text{constant}} q_k^{c+\beta_1-1} (1 - q_k)^{C-c+\beta_2-1} \mathrm{d}q_k$$

$$= \binom{C}{c} \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)} \int_0^1 {q_k}^{c+\beta_1-1} (1 - q_k)^{C-c+\beta_2-1} \mathrm{d}q_k .$$

For Beta Distribution is True

by Bronstein & Semendjajew (1983)

$$\int_0^1 q^{b_1-1}(1-q)^{b_2-1}\mathrm{d}q = \frac{\Gamma(b_1)\Gamma(b_2)}{\Gamma(b_1+b_2)} \qquad \forall b_1, b_2 > 0.$$

For $b_1 = c + \beta_1,\ b_2 = C - c + \beta_2$ and $\beta_1,\ \beta_2 > 0$ it results:

$$P(Y=c) = \binom{C}{c}\frac{\Gamma(\beta_1+\beta_2)\Gamma(c+\beta_1)\Gamma(C-c+\beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)\Gamma(C+\beta_1+\beta_2)}$$

With the moments of distribution

$$E(Y) = \frac{C\beta_1}{\beta_1+\beta_2} \qquad \text{(expectation value)}$$

and

$$Var(Y) = \frac{C\beta_1\beta_2(\beta_1+\beta_2+C)}{(\beta_1+\beta_2)^2(\beta_1+\beta_2+1)} \qquad \text{(variance).}$$

Beta Binomial Model for Heterogeneous Populations

$$P(Z = q|Y = c) = \frac{\binom{C}{c} q^c (1-q)^{C-c} \cdot f(q|.)}{P(Y = c)}$$

With the related conditional expectation value:

$$\begin{aligned}
E(Z = q|Y = c) &= \int_0^1 q \cdot P(Z = q|Y = c)\,\mathrm{d}q \\
&= \cdots \\
&= \underbrace{\frac{\beta_1}{C + \beta_1 + \beta_2}}_{a} + \underbrace{\frac{1}{C + \beta_1 + \beta_2}}_{b} c =: \hat{q}.
\end{aligned}$$

Unregulated Purchase Intention

With ρ $(0 \leq \rho \leq 1)$ as probability that the purchase intention changes:

$$\hat{p}_u = \rho \hat{q}^{new} + (1-\rho)\hat{q} = \rho \underbrace{\frac{\beta_1}{\beta_1 + \beta_2}}_{\substack{\text{expectation value} \\ \text{of beta distribution}}} + (1-\rho)\hat{q}.$$

Δ is the systematic bias of the unregulated purchase intention towards the observed purchase probabilities.

$$\hat{p} = \hat{p}_u - \Delta.$$

MODEL CALIBRATION

- Methods of moments
  (→ upcoming example)
- Restricted maximum likelihood method

$$\hat{m}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

$$\theta_1 = \mu = E(X)$$

$$\theta_2 = \sigma^2 = E(X^2) - E(X)^2$$

Example for Long-Living Consumer Goods

- Five categories of purchase intention and their correspondent purchase frequencies

	Caravan		Colour TV		F./D. Freeze		Spin dryer		Tape Rec.	
c	n_c	n_c^{Buy}	n_c	n_c^{Buy}	n_c	n_c^{Buy}	n_c	n_c^{Buy}	n_c	n_c^{Buy}
0	263	3	238	11	242	4	233	7	239	13
1	9	1	21	0	22	0	28	1	25	2
2	3	0	12	3	10	1	7	1	9	0
3	1	0	2	1	0	0	8	2	3	1
4	0	0	3	1	2	0	0	0	0	0

- Estimated purchase probabilities and purchase frequencies

	Caravan		Colour TV		F./D. Freeze		Spin dryer		Tape Rec.	
c	$\hat{p}$	$\hat{n}_c^{Buy}$	$\hat{p}$	$\hat{n}_c^{Buy}$	$\hat{p}$	$\hat{n}_c^{Buy}$	$\hat{p}$	$\hat{n}_c^{Buy}$	$\hat{p}$	$\hat{n}_c^{Buy}$
0	0,016	4,21	0,048	11,42	0,026	6,29	0,034	7,92	0,059	14,10
1	0,042	0,38	0,119	2,50	0,048	1,06	0,071	1,99	0,077	1,93
2	0,070	0,21	0,190	2,28	0,070	0,70	0,108	0,76	0,096	0,86
3	0,092	0,09	0,262	0,52	0,093	0,00	0,146	1,17	0,115	0,35
4	0,118	0,00	0,333	1,00	0,115	0,23	0,183	0,00	0,134	0,00

Structural Aspects of Purchase Intention Models

General Problem

- What makes a “good model”?
- How can models be more accepted?
- Where a the limitations?

Classification of Marketing Models

- Micro ↔ macro
- Data based ↔ theory based
- Descriptive ↔ explanatory ↔ forecast
- Behavioral ↔ statistical
- Generalized ↔ ad hoc
- Static ↔ dynamic
- Qualitative ↔ quantitative

Source: Williams, 1992.

Criteria for the Evaluation of Marketing Models

- Informational value
- Simplicity
- Flexibility
- Check ability
- Originality
- Transparency
- Robustness
- Validity
- Completeness

Homo Oeconomicus

A Well-Known Marketing Model

- Complete knowledge of his needs.
- Complete market review.
- No brand preferences.
- Unlimited information storage capacity.
- Is not subject to influences by the social environment.
- Alternatives:
    - “Homo Sociologicus“
    - “Satisfier“

Interpreting Meaning

Interpretation refers to the meaning that people assign to phenomena, either stimuli from the outside world or ideas and concepts from the person's own mind.

- **Priming:** where consumers assign meaning based on the set of held beliefs.
- **Symbolic consumption:** where meanings are attached to the act of consuming the goods, for example trendiness, wealth, femininity, etc.
- **Stimulus ambiguity:** where consumers project their own experiences and aspirations to assign meaning.
- **Stimulus organization:** where people relate incoming sensations to imagery of other sensations already existing in memory, based on fundamental organization principles.

Source: Solomon et al., 2006.

The Role of Symbolism in Interpretation

Some marketers are turning to semiotics to understand how consumers interpret the meanings of symbols.

- **Semiotics** examines the correspondence between signs and symbols and their role in the assignment of meaning.
- **Products** are given meanings by their producers and we rely on advertising to work out what those meanings are. Advertising serves as a kind of culture/consumption dictionary.
- **Hyper reality** - refers to the becoming real of what is initially 'hype'.

Source: Solomon et al., 2006.

Excursus: Multinomial Logit Model

- $P(Y_k = m)$ choice of brand m by customer k is calculated as follows:

$$
\begin{aligned}
P(Y_k = m) &= P(U_{km} \geq U_{kj}, \quad \forall j \neq m) \\
&= P(u_m + \epsilon_{km} \geq u_j + \epsilon_{kj}, \quad \forall j \neq m) \\
&= P(u_m - u_j + \epsilon_{km} \geq \epsilon_{kj}, \quad \forall j \neq m) \qquad \forall k, m
\end{aligned}
$$

- Common multivariat distribution $F(\epsilon_{k1},\ldots, \epsilon_{kM})$ with density function $f(\epsilon_{k1}, \ldots, \epsilon_{kM})$ of residuals

$$
P(Y_k = m) = \int\limits_{-\infty}^{u_m - u_1 + \epsilon_{km}} \cdots \int\limits_{-\infty}^{u_m - u_M + \epsilon_{km}} f(\epsilon_{k1}, \cdots, \epsilon_{km}, \cdots, \epsilon_{kM}) \, d\epsilon_{kM} \cdots d\epsilon_{k1}
$$

Excursus: Multinomial Logit Model (cont.)

- Stochastical independence of residuals

$$P(Y_k = m) = \int_{-\infty}^{u_m - u_1 + \epsilon_{km}} \cdots \int_{-\infty}^{u_m - u_M + \epsilon_{km}} \prod_{\substack{j=1 \\ j \neq m}}^{M} f(\epsilon_{kj})\, d\epsilon_{kj} \cdot \int_{-\infty}^{\infty} f(\epsilon_{km})\, d\epsilon_{km}$$

$$= \prod_{\substack{j=1 \\ j \neq m}}^{M} F(\epsilon_{k1}, \ldots, \epsilon_{k,m-1}, \epsilon_{k,m+1}, \ldots, \epsilon_{kM}) \cdot \int_{-\infty}^{\infty} f(\epsilon_{km})\, d\epsilon_{km} \qquad \forall\, k, m$$

- Utility of chance equates distribution of extreme values

$$P(Y_k = m) = \prod_{\substack{j=1 \\ j \neq m}}^{M} \exp(-\exp(-u_m + u_j - \epsilon_{km})) \cdot \int_{-\infty}^{\infty} \exp(-\epsilon_{km} - \exp(-\epsilon_{km}))\, d\epsilon_{km}$$

$$= \int_{-\infty}^{\infty} \exp\left(\sum_{\substack{j=1 \\ j \neq m}}^{M} (-\exp(-u_m + u_j - \epsilon_{km})) - \epsilon_{km} - \exp(-\epsilon_{km}) \right) d\epsilon_{km}$$

EXCURSUS: MULTINOMIAL LOGIT MODEL (CONT.)

- Continuation

$$= \int_{-\infty}^{\infty} \exp\left(-\exp(-\epsilon_{km}) \cdot \sum_{\substack{j=1 \\ j\neq m}}^{M} \exp(u_j - u_m) - \epsilon_{km} - \exp(-\epsilon_{km}) \right) \, d\epsilon_{km}$$

$$= \int_{-\infty}^{\infty} \exp\left(-\epsilon_{km} - \exp(-\epsilon_{km}) \cdot \left(1 + \sum_{\substack{j=1 \\ j\neq m}}^{M} \frac{\exp(u_j)}{\exp(u_m)} \right) \right) \, d\epsilon_{km} \qquad \forall\, k, m$$

- Simplification

$$Q_m := \ln\left(1 + \sum_{\substack{j=1 \\ j\neq m}}^{M} \frac{\exp(u_j)}{\exp(u_m)} \right) = \ln\left(\sum_{j=1}^{M} \frac{\exp(u_j)}{\exp(u_m)} \right) \qquad \forall\, m$$

EXCURSUS: MULTINOMIAL LOGIT MODEL (CONT.)

- Thus:

$$P(Y_k = m) = \int_{-\infty}^{\infty} \exp\left(-\epsilon_{km} - \exp(-\epsilon_{km}) \cdot \exp(Q_m)\right) \, d\epsilon_{km}$$

$$= \int_{-\infty}^{\infty} \exp\left(-\epsilon_{km} - \exp(-\epsilon_{km} + Q_m)\right) \, d\epsilon_{km} \qquad \forall\, k, m$$

- Substitution $\tilde{\epsilon}_{km} := \epsilon_{km} - Q_m$

$$P(Y_k = m) = \exp(-Q_m) \cdot \int_{-\infty}^{\infty} \exp\left(-\epsilon_{km} + Q_m - \exp(-\epsilon_{km} + Q_m)\right) \, d\epsilon_{km}$$

$$= \exp(-Q_m) \cdot \int_{-\infty}^{\infty} \underbrace{\exp\left(-\tilde{\epsilon}_{km} - \exp(-\tilde{\epsilon}_{km})\right)}_{\substack{\text{density function} \\ \text{extreme values}}} \, d\tilde{\epsilon}_{km} = \exp(-Q_m) \cdot 1$$

EXCURSUS: MULTINOMIAL LOGIT MODEL (CONT.)

- Continuation

$$= \frac{1}{\sum_{j=1}^{M} \frac{\exp(u_j)}{\exp(u_m)}} = \frac{1}{\frac{1}{\exp(u_m)} \sum_{j=1}^{M} \exp(u_j)} = \frac{\exp(u_m)}{\sum_{j=1}^{M} \exp(u_j)}$$

Availability

- Estimation of all customers → availability of all customer

$$P(Y = m) = \frac{\exp(u_m)}{\sum_{j=1}^{M} \exp(u_j)} =: p_m \qquad \forall\, m$$

EXCURSUS: MULTINOMIAL LOGIT MODEL (CONT.)

Linear Analysis von utility Terms

- For all customers

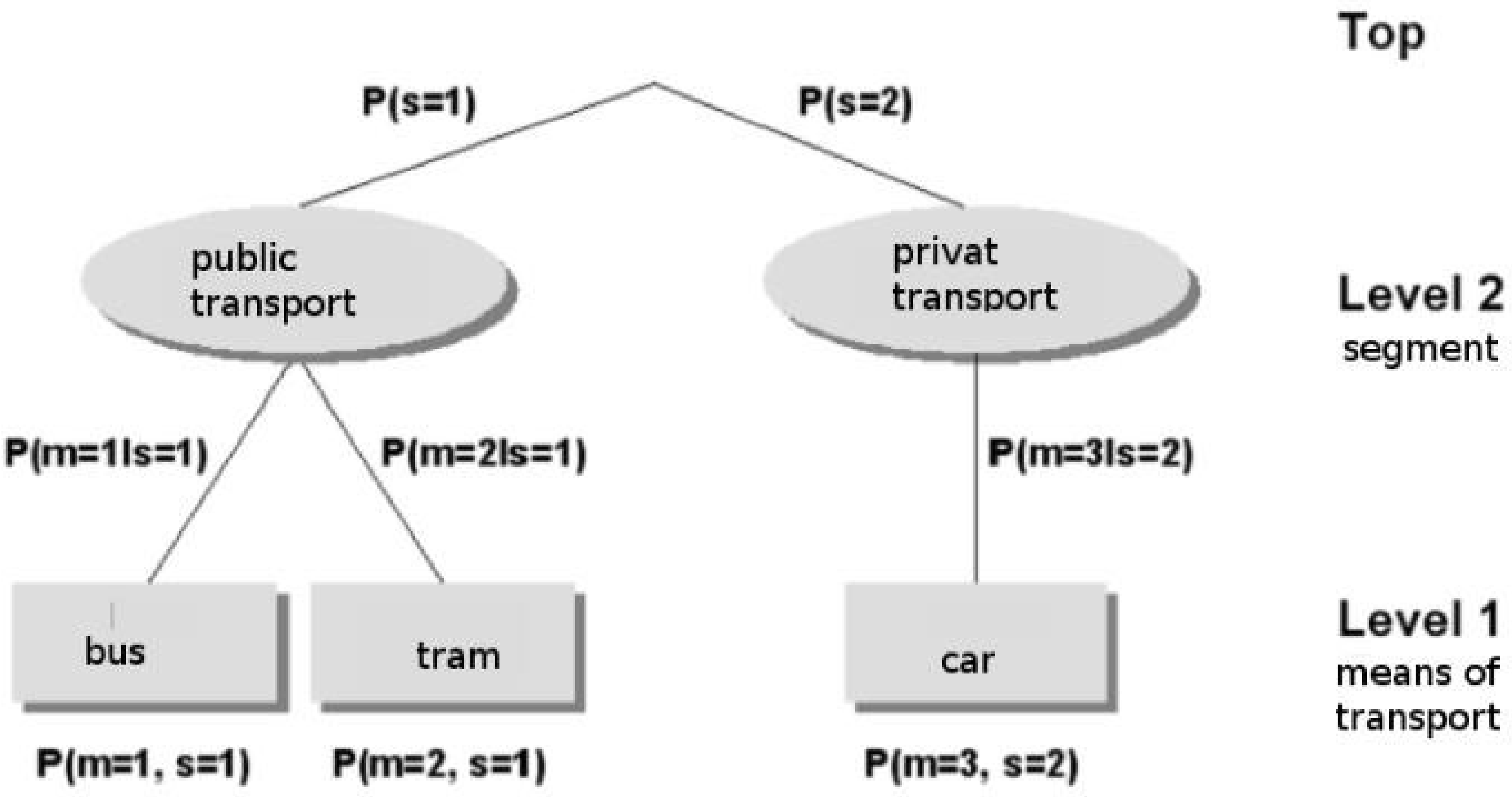
$$u_m = \eta_m + \tilde{\eta}_1 x_{m1} + \cdots + \tilde{\eta}_L x_{mL} \qquad \forall\ m$$

Elasticities

$$ELA_{p_{mt}, x_{jlt}} = \frac{\partial p_{mt}}{\partial x_{jlt}} \cdot \frac{x_{jlt}}{p_{mt}} = \begin{cases} \eta_{ml} x_{mlt} (1 - p_{mt}) & \text{für } m = j \\ -\eta_{jl} x_{jlt} p_{jt} & \text{für } m \neq j \end{cases} \qquad \forall\ m, j, l, t$$

EXCURSUS: MULTINOMIAL LOGIT MODEL (CONT.)

Idea Nested Logit



3.	THEORY BUILDING AND HYPOTHESES
3.1	THEORY BUILDING
3.2	HYPOTHESES GENERATION

THE RESEARCH DESIGN

Problem Definition

↓

Hypothesis Generation

↓

Decision on the Type(s) of Study appropriate to the Problem

↓

Decision on the Method of Data Gathering

↓

Development of an Analysis Plan

↓

Data Collection

↓

Performance of the Analysis

↓

Drawing Conclusions and Recommendations

Source: FAO, n.d.

Hypotheses Development

- Research questions are too broad to be directly testable.
- Hypotheses make the research question more precise.
- Use of three main sources to develop hypotheses:
  1. Draw on previous research efforts (exploratory research),
  2. Theory from other disciplines (e.g., psychology, sociology, marketing, or economics),
  3. Management experience with related problems, knowledge of the problem situation and the use of judgment.

Source: FAO, n.d.; Aaker et al., 2007, p. 57 ff.

Hypothesis – Definition

"An unproven statement or proposition about a factor or phenomenon that is of interest to the researcher. It may, for example, be a tentative statement about relationships between two ore more variables as stipulated by the theoretical framework or the analytical model. Often a hypothesis is a possible answer to the research question."

(Malhotra, 2007, p. 53)

"A hypothesis is a logically conjectured relationship between two or more variables expressed in the form of testable statements."

(Sekaran & Bougie, 2010, p. 87)

Good Research Hypothesis

A **good** hypothesis has the following three characteristics:

- **It is theoretically grounded:** It is based upon literature relevant to the topic.
- **It specifies the relationship between the values of two or more variables:** This includes both the connection and tendency. E.g., frequency of pub visits is negatively associated with studying for university.
- **It makes a testable comparison using empirical data:** The collected data can disprove the hypothesis. For instance, it is possible in the previous example for frequency of pub visits is in fact positively associated with studying for university.

Source: Based on Palazzolo & Roberts, 2010.

Criteria for a good Research Hypothesis

- A good hypothesis is stated in **declarative form** and **not** as a **question**.

  "Are tea drinker healthier than coffee drinker?" is not declarative, but "Tea drinker are healthier than coffee drinker" is.

- A good hypothesis **posits** an **expected relationship** between variables and **clearly states a relationship** between variables.

  "Students who participate in a workshop for generating research hypothesis three hours per week will score higher on a test of hypothesis comprehension than students who do not."
  → Clear relationship between hours of workshop and test score.

Source: Based on Sage Publication, n.d.

CRITERIA FOR A GOOD RESEARCH HYPOTHESIS (CONT.)

- Hypotheses **reflect** the **theory** or **literature** on which they are **based**. A good hypothesis has a **substantive link** to existing literature and theory.

    In the above example, let's assume there is literature indicating that reading to students is one way to increase their comprehension. The hypothesis is a test of that idea.

- A hypothesis should be **brief** and **to the point**.
- Good hypotheses are **testable** hypotheses: One can actually carry out the intent of the question reflected by the hypothesis.

    E.g., number of hours of student studying and outcome scores as measured by a test of comprehension are all objective and can be incorporated reliably.

Source: Based on Sage Publication, n.d.

Criteria for a good Research Hypothesis

Conclusion

Finally, a good research hypothesis **combines** all of the above to be **understandable** and **easy to envision** how it fits into the larger world of the research question. After reading such a hypothesis, the reader should have a **good grasp** of which **direction the research** is taking and what some of the **implications** for its **testing** might be.

Source: Sage Publication, n.d.

Research Hypothesis

- A research hypothesis is usually expressed using an **if-then** structure.
  E.g., "***If*** aspirin is taken ***then*** pain will be reduced."
- But not all **if-then** statements are hypotheses.
  E.g., "If I play the lottery, then I will get rich." This is a simple **prediction**.
- In a formalized hypothesis, a **tentative relationship** is stated. E.g., If the frequency of winning is **related** to frequency of buying lottery tickets. "**Then**" is followed by a **prediction** of **what will happen** if you increase or decrease the frequency of buying lottery tickets.

Source: Access Excellence, n.d.; Bortz & Döring, 2006, p. 4 ff.; University of West England, 2007.

RESEARCH HYPOTHESIS (CONT.)

- If you always ask yourself that if **one thing** is **related to another**, then you should be able to **test** it.
- A research hypothesis may also resemble these:
  - Gender has no effect on intelligence.
  - Drug X has an effect on flu symptoms.
  - There is more sunshine in December than January.
- These are all statements which can be tested through a data collection strategy: **The experiment**.

Source: Access Excellence, n.d.; Bortz & Döring, 2006, p. 4 ff.; University of West England, 2007.

EXAMPLES OF BAD RESEARCH HYPOTHESES

- **Untestable:**
  "The more supportive of political authorities a child is, the less likely the child will be to engage in political dissent as an adult."
  No data!
- **Not specific:**
  A country's geographical location matters for the type of political system it develops.
  Better: "The more borders a country shares with democratic nations, the more likely it is to have a democratic system."
- **Non-directional:**
  "Age determines whether a person is tolerant of social protest."
  Better: "Older people are less tolerant of social protest than younger people."

Source: Kapanadze, n.d.

Examples of Bad Research Hypotheses

- **Implausible:**

  "People who eat dry cereal for breakfast are more likely to be liberal than people who eat eggs."
- **Narrow:**

  "The united states has more murders than other countries because so many people own guns there."

  Better: "Countries with more guns per capita will experience more murders than countries with fewer guns."
- **Tautology:**

  "The less support there is for a country's political institutions, the more tenuous the stability of that country's political system."

Source: Kapanadze, n.d.

Characteristic of Hypotheses

- Two key characteristics which all hypotheses must have:
    1. Statements of the relationship between variables.
    2. Carry clear implications for testing the stated relations.
- Imply that it is relationships, rather than variables, which are tested.
- Hypotheses specify how the variables are related and that these are measurable or potentially measurable.
- Hypotheses are falsifiable.

→ **Statements lacking any or all of these characteristics are not research hypotheses.**

Source: FAO, n.d.

EXAMPLE: HYPOTHESIS I

- **"Red meat consumption increases as real disposable incomes increase."**

- What do you think – have the criteria for a hypothesis been met?

red meat consumption | **disposable incomes**

Variable 1 | Variable 2

measurable

testable

Source: FAO, n.d.

Example: Hypothesis II

- **"There is no relationship between a farmer's educational level and his degree of innovativeness with respect to new farming technologies."**

- What do you think – have the criteria for a hypothesis been met?

- Clear statement of the relationship
- Measurability?
  - **"[…] a farmer's educational level "**
  - **"level of education"**

Resolve to have a hypothesis!

Source: FAO, n.d.

DMCC DIALOG MARKETING COMPETENCE CENTER

Hypothesis Tableau

- A hypothesis tableau is a systematization of items and hypothesis.
- Shows if relevant survey items are covered by at least one question.
- Empty row refers to a missing operationalization.
- If one question does not correspond with at least one hypothesis, it is appropriate to verify whether the elicitation of this aspect is necessary.

Source: Decker & Wagner, 2002, p. 164.

Section of a Hypothesis Tableau

	Question 1	Question 2	Question 3	Question 4	Question 5	…
Hypothesis 1		×		×		…
Hypothesis 2	×		×			…
Hypothesis 3		×			×	…
Hypothesis 4		×				…
Hypothesis 5	×					…
⋮	⋮	⋮	⋮	⋮	⋮	⋱

Source: Decker & Wagner, 2002, p. 164.

3. Theory Building and Hypotheses

3.1 Theory Building

3.2 Hypotheses Generation

3.3 Hypotheses Testing

A General Procedure for Hypothesis Testing

1. Formulate H_0 and H_1 (translate content to statistical hypothesis).
2. Select an appropriate test.
3. Choose the level of significance, α.
4. Collect data and calculate the test statistic.
5. Determine the probability associated with the test statistic under the H_0, using the sampling distribution of test statistic. Alternatively, determine whether the test statistic that divide the rejection and non-rejection regions.
6. Compare probability associated with the test statistic with the level of significance specified. Alternatively, determine whether the test statistic has fallen into the rejection or the non-rejection region.
7. Reject or do not reject H_0.
8. Draw a (marketing) research conclusion.

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 1

- A **null hypothesis** (H_0) is a statement of the status quo, one of no difference or no effect.
- If H_0 is not-rejected, no changes will be made.
- An **alternative hypothesis** (H_1) is one in which some differences or effect are expected.
- Accepting H_1 will lead to changes in options/actions.
- H_1 is the opposite of H_0.
- H_0 is always the hypothesis that is tested:

  Refers to a specific value of the population parameter (e.g., μ, σ, π), not a sample statistic (e.g., $\bar{x}$).
- One or two outcomes of statistical test:
  - $H_0 \rightarrow$ rejected; $H_1 \rightarrow$ accepted
  - $H_0 \rightarrow$ not-rejected based on evidence

Source: Malhotra, 2007, p. 464 ff.

EXERCISE: HYPOTHESES BUILDING

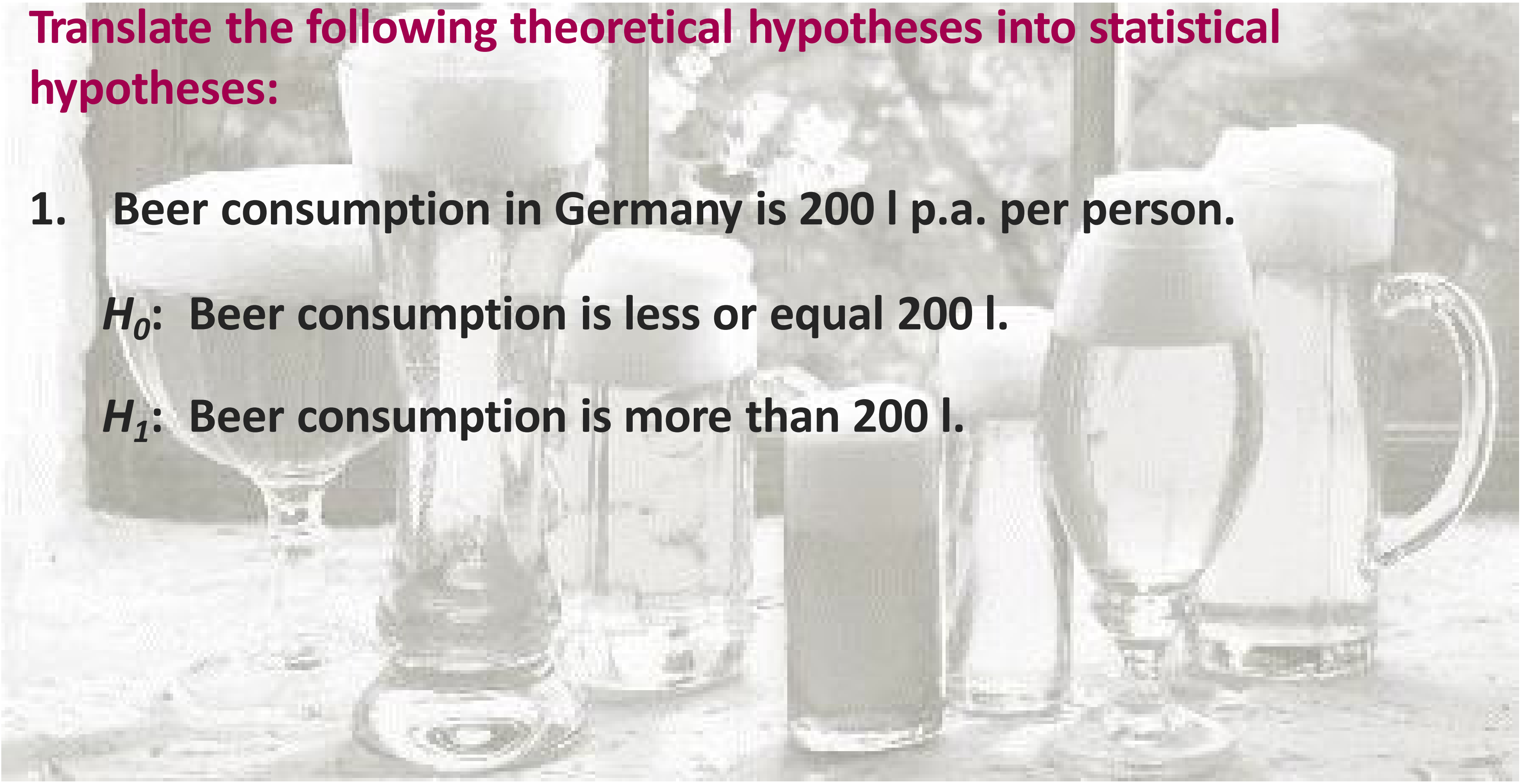
Translate the following theoretical hypotheses into statistical hypotheses:

1. Beer consumption in Germany is 200 l p.a. per person.
2. The consumption of chocolate biscuits increases with the consumption of coffee.
3. People tend to drink more tea in winter than in summer.



EXERCISE: HYPOTHESES BUILDING

SOLUTION: THEORETICAL HYPOTHESIS 1



Translate the following theoretical hypotheses into statistical hypotheses:

1. **Beer consumption in Germany is 200 l p.a. per person.**

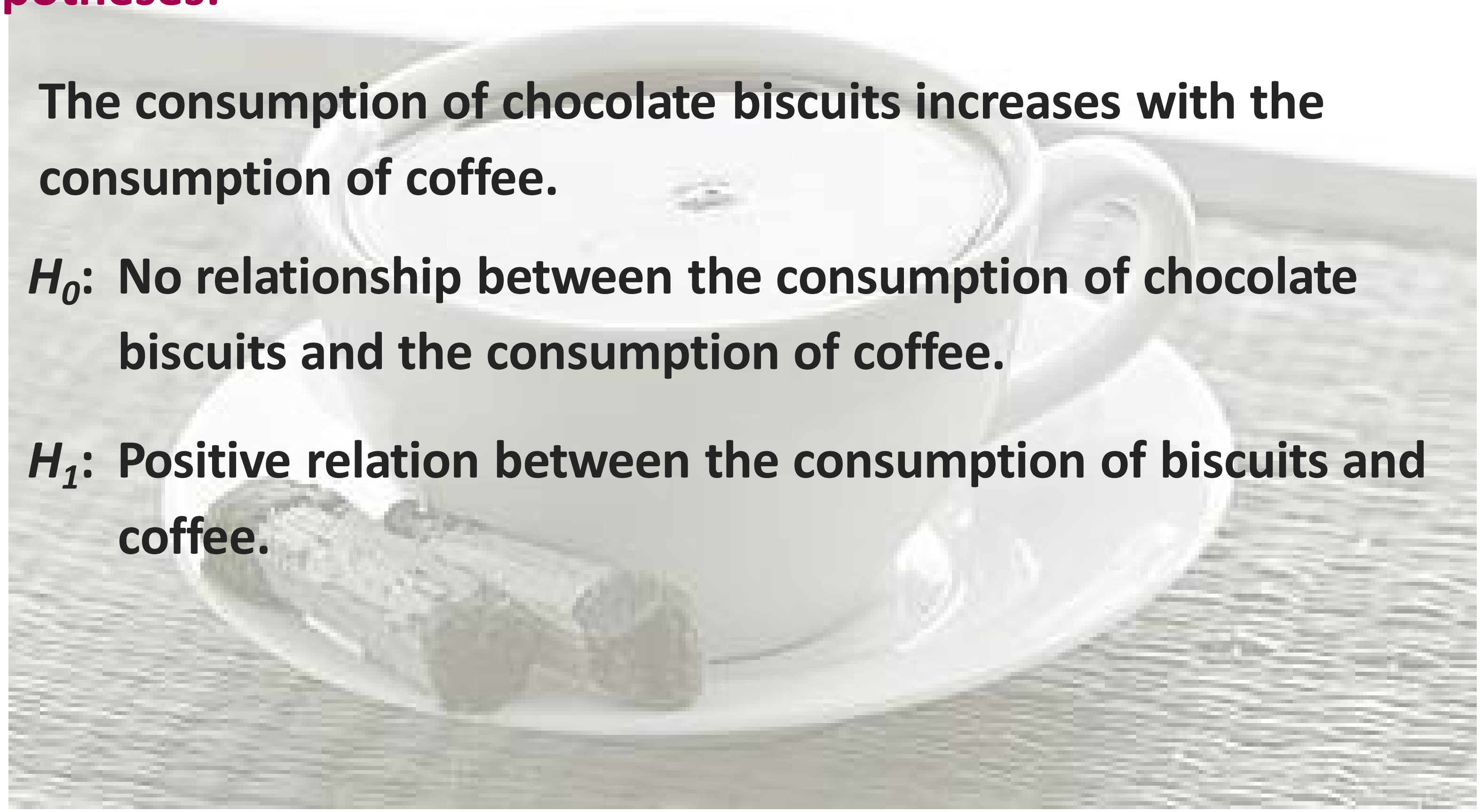
   H_0: Beer consumption is less or equal 200 l.

   H_1: Beer consumption is more than 200 l.

Exercise: Hypotheses Building

Solution: Theoretical Hypothesis 2

Translate the following theoretical hypotheses into statistical hypotheses:



2. **The consumption of chocolate biscuits increases with the consumption of coffee.**

   ***H_0*: No relationship between the consumption of chocolate biscuits and the consumption of coffee.**

   ***H_1*: Positive relation between the consumption of biscuits and coffee.**

DMCC DIALOG MARKETING COMPETENCE CENTER

Exercise: Hypotheses Building

Solution: Theoretical Hypothesis 3

Translate the following theoretical hypotheses into statistical hypotheses:

3. People tend to drink more tea in winter than in summer.

H_0: Consumption of tea is equal in winter and summer.

H_1: People significantly drink more tea in winter than in summer.

A General Procedure for Hypothesis Testing

Step 1 (cont.)

- *Example:*

  A department store considers the introduction of an online shopping service. The new service will be introduced if more than 40% of the internet users shop online.

- *Hypothesis formulation:*

  H_0: $\pi \leq 0.40$

  H_1: $\pi > 0.40$

- If H_0 is rejected, then H_1 will be accepted and the new online shopping service will be introduced. If H_0 is not rejected, the new service shouldn‘t be introduced unless additional evidence is obtained.

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 1 (cont.)

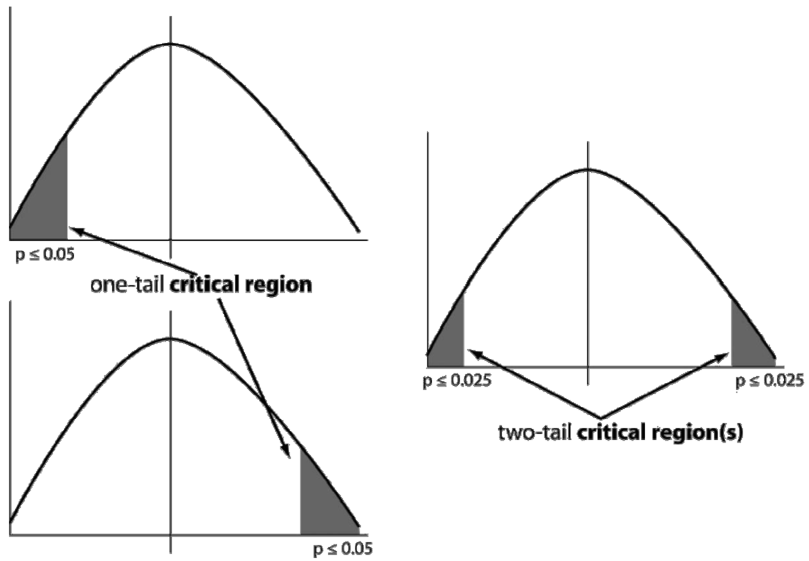
H_1 is expressed directionally

→ **One-tailed test:** The proportion of internet users who use the internet for shopping is greater than 0.40. On the other hand one can n also determine whether the proportion of internet users who shop via the internet is different from 40%.

→ **Two-tailed test:**

H_0: $\pi = 0.40$

H_1: $\pi \neq 0.40$



Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing
Step 2

- The test statistic measures how close the sample has come to the H_0.
- Often follows a well known distribution, such as the normal, *t*, or chi-square distribution.
- For the example the *z** statistic, which follows the standard normal distribution, is appropriate:

$$z = \frac{p - \pi}{\sigma_p}$$

where

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

* *z* test is used when σ is known, *t* test is used when σ is unknown.

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 3

Risk that an incorrect conclusion will be reached, when drawing inferences about a population → two types of errors:

- **Type I Error** (α-error)
  - Occurs when the sample results lead to the rejection of a null hypothesis that is in fact true.
  - *Example:* Type I error would occur if the conclusion was that the proportion of customers preferring the new service was > 0.40, when in fact is was ≤ 0.40.
  - Also called the **level of significance**.

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 3 (cont.)

- **Type II Error** (β-error)
  - Occurs when the sample results lead to the non-rejection of a null hypothesis that is in fact false.
  - *Example:* Type II error would occur if the conclusion was that the proportion of customers preferring the new service was ≤ 0.40, when in fact it was > 0.40.
  - Probability of Type II error is denoted by β.
  - Magnitude of β depends on the actual value of the population parameter.
  - The complement (1 – β) of the probability of a Type II error is called the **power of a statistical test**.

Source: Malhotra, 2007, p. 464 ff.

OPERATING CHARACTERISTIC (OC)

- Shows first hints with regard to suitable sample size.
- Defined as probability:

$$\beta = f(\Theta_1)$$

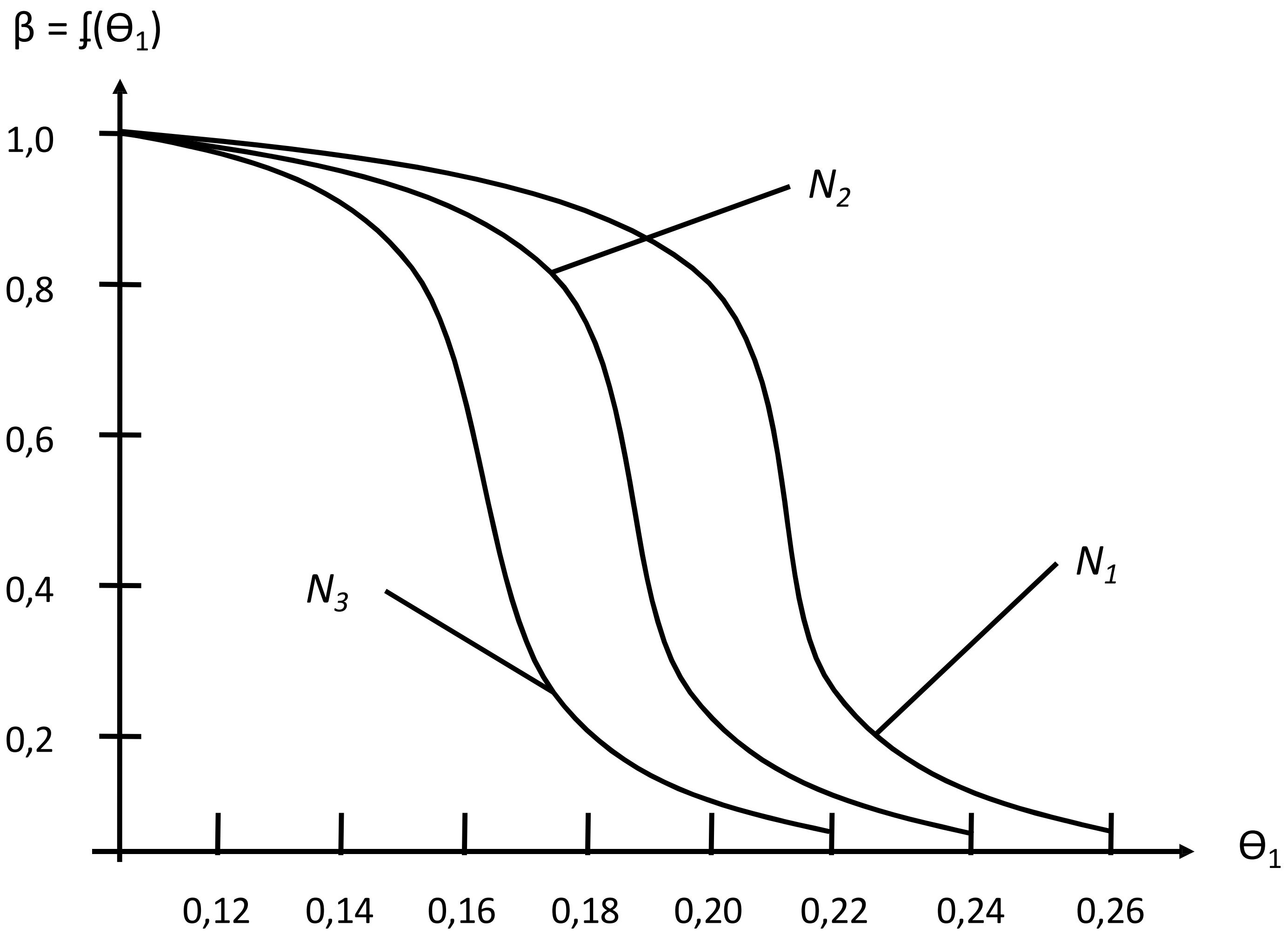
- If H_1 is valid, then Θ_1 is a true parameter.
- The probability

$$1 - \beta = 1 - f(\Theta_1)$$

describes the **power of a test**.

Source: Decker & Wagner, 2002, p. 105 f.

OC-CURVES



Source: Decker & Wagner, 2002, p. 106.

OC-CURVES – DESCRIPTION

- Varying sample size $N_1 < N_2 < N_3$ with identical level of significance α.
- The larger the sample size, the steeper the OC-curve.
- **Selectivity** of the test increases with increasing of "steepness".
- The probability of an incorrect non-rejection of e.g., H_1: $\Theta = \Theta_1 = 0{,}18$ would be smaller with sample size N_2 than with N_1 and larger than with N_3.

Source: Decker & Wagner, 2002, p. 106.

A GENERAL PROCEDURE FOR HYPOTHESIS TESTING

STEP 3 (CONT.)

- **Power of a Test**
  - The probability (1 – β) of rejecting the null hypothesis when it is in fact false and should be rejected.
  - Although β is unknown, it is related to α.
  - It is necessary to balance the two types of errors.
  - α is often set at 0.05 (0.01).
  - The level of α will determine the level of β.
  - By increasing the sample size the risk of α and β can be controlled.

    → A given level of α, increasing the sample size will decrease β.

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 4

- After taking the desired α and β errors into account the sample size is determined.
- Required data are collected.
- Value of test statistic computed.
- *Example:* 30 users were surveyed; 17 indicated that they use internet for shopping → value of sample proportion:

  $P = 17/30 = 0.567$

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 4 (cont.)

- The value of σ_p can be determined as follows:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

$$= \sqrt{\frac{(0.40)(0.60)}{30}}$$

$$= 0.089$$

- The test statistic *z* can be calculated as follows:

$$z = \frac{p-\pi}{\sigma_p}$$

$$= \frac{0.567-0.40}{0.089}$$

$$= 1.88$$

Source: Malhotra, 2007, p. 464 ff.

Critical Values for Normal Distribution

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

Source: Kosfeld, 2012a, p. 1.

A General Procedure for Hypothesis Testing

Step 5

- Using standard normal tables, the probability of obtaining a z value of 1.88 can be calculated.
- Shaded are between $-\infty$ and 1.88 is 0.9699.
- The area to the right of z = 1.88 is 1.0000 – 0.9699 = 0.0301.
- Alternatively: critical value of z, which will give an area to the right side of the critical value of 0.05, is between 1.64 and 1.65 and equals 1.645.
- Note: In determining the critical value of test statistic, the area to the right of critical value is either α (one-tail test) or $\alpha/2$ (two-tail test).

Source: Malhotra, 2007, p. 464 ff.

A General Procedure for Hypothesis Testing

Step 6 & 7

- Probability associated with calculated/observed value of test statistic is 0.0301 → probability of getting a *p*-value of 0.567 when π = 0.40 → less than level of significance 0.05 → H_0 rejected.
- Alternative: z = 1.88 lies in rejection region, beyond the value of 1.645 → reject H_0.
- If probability associated with the calculated/observed value of test statistic is less than the level of significance (α) → H_0 rejected.
- If calculated value of test statistic is greater than the critical value of test statistic → H_0 rejected.

Source: Malhotra, 2007, p. 464 ff.

Basic Concepts of Hypothesis Testing

One Sample Test

Null hypothesis (H_0)	An assertion that a parameter in a statistical model takes a particular value, and is assumed true until experimental evidence suggests otherwise.
Alternative hypothesis (H_1)	Expresses the way in which the value of a parameter may deviate from that specified in the H_0, and is assumed true when the experimental evidence suggests that the H_0 is false.
Type 1 error	Rejecting the H_0 when it is, in fact, true.
Type 2 error	Accepting the H_0 when it is, in fact, false.
Test statistic	A function of sample of observations which provides a basis for testing the validity of the H_0.
Critical region	The H_0 is rejected when a calculated value of the test statistic lies within this region.
Critical value	The value which determines the boundary of the critical region.

Source: Centre for Innovation in Mathematics Teaching, n.d.

Basic Concepts of Hypothesis Testing (cont.)

One Sample Test

Significance level (α)	The size of the critical region; the probability of a Type 1 error.
One-tailed test	The critical region is located wholly at one end of the sampling distribution of the test statistic; H_1 involves $<$ or $>$ but not both.
Two-tailed test	The critical region comprises areas at both ends of the sampling distribution of the test statistic; H_1 involves $\neq$.

Source: Centre for Innovation in Mathematics Teaching, n.d.

Bonferroni Correction

- Method to address the problem of multiple comparisons.
- Based on the idea that if an experimenter is testing *n* dependent or independent hypotheses on a set of data, then one way of maintaining the family wise error rate is to test each individual hypothesis at a statistical significance level of *1/n* times what it would be if only one hypothesis were tested.
- If one wants the significance level for the whole family of tests to be α, then the **Bonferroni correction** would be to test each of the individual tests at a significance level of α/n.

Source: Brett et al., 2003, p. 3 ff.

BONFERRONI CORRECTION - EXAMPLE

Suppose we have *m* tests and each is designed to guarantee $P_{FA} \leq \alpha$. Then for any one test, the chance of a false alarm is α. However the prob of at least one false alarm among all the tests is much higher. This is the multiple testing problem.

Consider *m* hypothesis tests (a family of tests): $\mathbf{H}_{0i}$ v.s. $\mathbf{H}_{1i}$, where $i = 1, 2, ..., m$

The family-wise error rate ***(FWER)*** is the probability of one or more false alarms.

$$FWER = \mathbb{P}(\bigcup_{i=1}^{m} \{\text{decide } \mathbf{H}_{0i} \text{ when } \mathbf{H}_{1i} \text{ is true}\})$$

In many applications we want to control the ***FWER***.

Source: Nowak, 2010, p. 2.

BONFERRONI CORRECTION – EXAMPLE (CONT.)

$$\begin{array}{ccc} \mathbf{H}_{01} \text{ v.s. } \mathbf{H}_{11} & ; & \mathbf{H}_{02} \text{ v.s. } \mathbf{H}_{12} \\ t_1 \underset{H_{01}}{\overset{H_{11}}{\gtrless}} \gamma_1 & ; & t_1 \underset{H_{02}}{\overset{H_{12}}{\gtrless}} \gamma_1 \\ \mathbb{P}(t_1 > \gamma_1 \mid \mathbf{H}_{01}) \leq \alpha & ; & \mathbb{P}(t_2 > \gamma_2 \mid \mathbf{H}_{02}) \leq \alpha \end{array}$$

γ_1 and γ_2 are sets such that $P_{FA} \leq \alpha$ in both cases.

Recall that for any two events A and B, we have $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$ with equality *iff* $A \cup B = \emptyset$. This is called the union of events bound, union bound or Bonferroni's inequality.

Thus,

$$FWER = \mathbb{P}(t_1 > \gamma_1 \text{ or } t_2 > \gamma_2 \mid \mathbf{H}_{01}, \mathbf{H}_{02}) = \mathbb{P}(\{t_1 > \gamma_1 \mid \mathbf{H}_{01}\} \cup \{t_2 > \gamma_2 \mid \mathbf{H}_{02}\}) \leq 2\alpha$$

More generally, if we have m tests and each has an individual

$\mathbb{P}_{FA} \leq \alpha$, then $FWER \leq m\alpha$.

Source: Nowak, 2010, p. 2.

Bonferroni Correction – Example (cont.)

Definition: **Bonferroni Correction**

To guarantee a ***FWER*** for a family of m tests, we can set the $\mathbb{P}_{FA} \leq \frac{\alpha}{m}$ for each individual test.

Example: Threshold for multiple hypothesis testing

$$\mathbf{H}_{0i} : X_i \sim \mathcal{N}(0,1)$$
$$\mathbf{H}_{1i} : X_i \sim \mathcal{N}(\mu,1)$$
$$where\ i = 1, 2, ...m,\ \mu \geq 1$$

$$For\ the\ test:\ x_i \underset{H_{0i}}{\overset{H_{1i}}{\gtrless}} \gamma,\ if\ we\ use\ \gamma = Q^{-1}(\tfrac{\alpha}{m})\ instead\ of\ \gamma = Q^{-1}(\alpha),\ then\ FWER \leq \alpha$$

Source: Nowak, 2010, p. 2.

PARAMETRIC AND NONPARAMETRIC TESTS

Parametric Tests:

"Hypothesis-testing procedures that assume that the variables of interest are measured on at least an interval scale."

(Malhotra, 2007, p. 478)

Nonparametric Tests:

"Hypothesis-testing procedures that assume that the variables are measured on a nominal or ordinal scale."

(Malhotra, 2007, p. 478)

Hypothesis Tests related to Differences

Hypothesis Tests

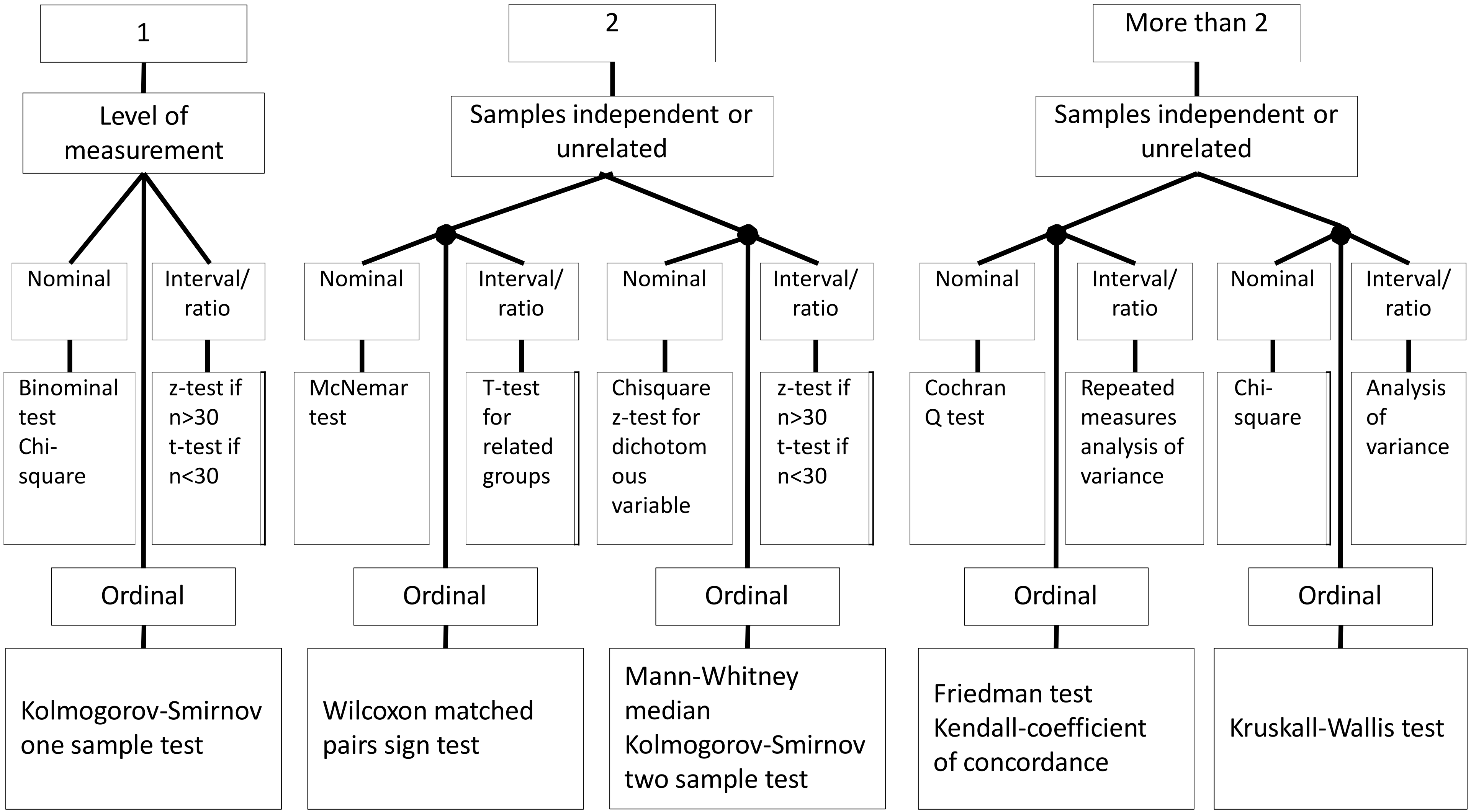
- Metric Data
  - One Sample
    - *t* Test
    - *z* Test
  - Two Samples
    - Independent Samples
      - Two-group *t* Test
      - *z* Test
    - Paired Samples
      - *Paired t* Test
- Non-metric Data
  - One Sample
    - Chi-Square
    - K-S*
    - Runs
    - Binominal
  - Two Samples
    - Independent Samples
      - Chi-Square
      - Mann-Whitney
      - Median
      - K-S
    - Paired Samples
      - Sign
      - Wilcoxon
      - McNemar
      - Chi-Square

* K-S = Kolmogorov-Smirnov Test

Source: Malhotra, 2007, p. 478.

4. SELECTING STATISTICAL TESTS

SELECTING STATISTICAL TESTS



4.

Selecting Statistical Tests

Parametric Tests

4.1

t Test

t Test

- ***t* test:** "A univariate hypothesis test using the *t* distribution, which is used when the standard deviation is unknown and the sample size is small." (Malhotra, 2007, p. 479)
- A ***t* test** is commonly used for making statements about the means of parent populations.
- Based on the Student's *t* statistic.
- "***t* statistic** assumes that the variable is normally distributed and the mean is known (or assumed to be known), and the population variance is estimated from the sample." (Malhotra, 2007, p. 479)
- ***t* distribution:** "a symmetric bell-shaped distribution that is useful for small sample ($n < 30$) testing." (Malhotra, 2007, p. 479)

Source: Malhotra, 2007, p. 479 f.

t Test – Example

A consumer group, concerned about the mean fat content of a certain grade of steakburger submits to an independent laboratory a random sample of 12 steakburgers for analysis.

The percentage of fat in each of the steakburger is as follows:

21	18	19	16	18	24	22	19	24	14	18	15

The manufacturer claims that the mean fat content of this grade is less than 20%.

Assuming percentage fat content to be normally distributed with an unknown standard deviation, carry out an appropriate hypothesis test in order to advise the consumer group as to the validity of the manufacturers claim.

Source: Centre for Innovation in Mathematics Teaching, n.d.

t TEST – EXAMPLE: SOLUTION

Solution:

Denoting the percentage of fat content by *X*, then *X* ~ *N*(μ, σ^2) then

$$\frac{\bar{x}-\mu}{\frac{\hat{\sigma}}{\sqrt{n}}}$$

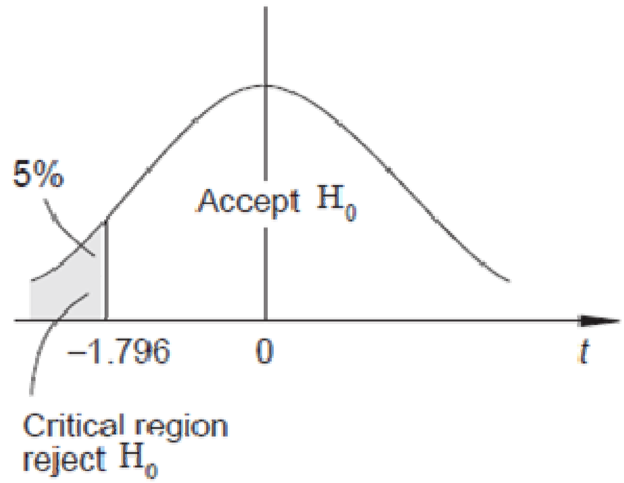
Is a *t* statistic with degrees of freedom *v* = *n-1*, where

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum(x-\bar{x})^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]$$

Source: Centre for Innovation in Mathematics Teaching, n.d.

t Test – Example (cont.)

H_0	$\mu = 20\%$
H_1	$\mu < 20\%$ (one-tailed)
Significance level	$\alpha = 0.05$
Degrees of freedom	$v = n\text{-}1 = 11$
Critical region	$t < -1.796$



Under H_0, $\overline{X} \sim N\left(20, \frac{\sigma^2}{12}\right)$

Source: Centre for Innovation in Mathematics Teaching, n.d.

DMCC DIALOG MARKETING COMPETENCE CENTER

t Test – Example: Solution

Now the test statistic is

$$t = \frac{\bar{x} - \mu}{\frac{\hat{\sigma}}{\sqrt{n}}}$$

with $\bar{x} = 19$ and $\hat{\sigma}^2 = \frac{1}{11}\left(4448 - \frac{228^2}{12}\right) = 10.545$ and so

$\hat{\sigma} = 3.25$

Hence $t = \frac{19 - 20}{\frac{3.25}{\sqrt{12}}} = -1.07$

This value does not lie in the critical region. Thus there is no evidence, at the 5% level of significance, to support the manufacturer's claim.

Source: Centre for Innovation in Mathematics Teaching, n.d.

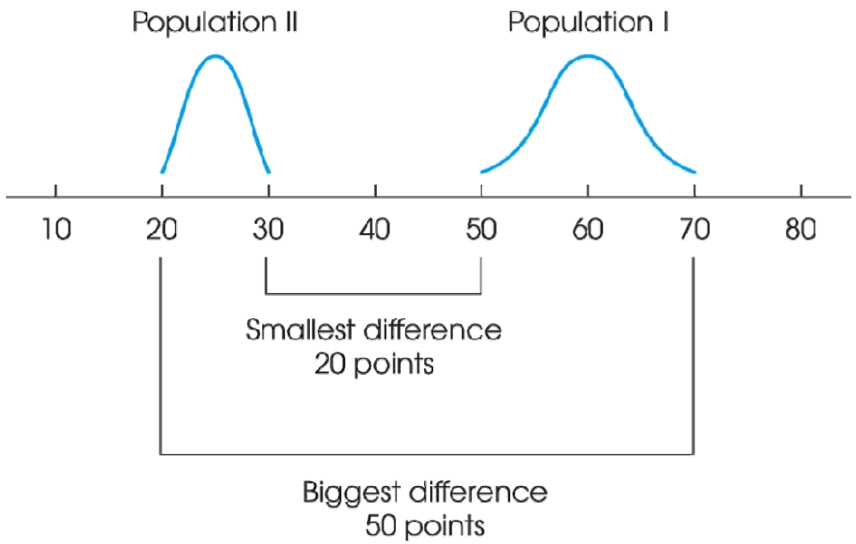
Homogeneity of Variance Assumption

- Assumption requires that the variance within each population is equal for all populations (two or more, depending on the method).
- Assumption is used in e.g., the two-sample *t* test.
- Assumption is necessary in order to justify pooling the two sample variances and using the pooled variance in the calculation of the *t* statistic.
- If the assumption is violated, then the *t* statistic contains two questionable values:
  - (1) The value for the population mean difference which comes from the null hypothesis.
  - (2) The value for the pooled variance.

Source: Mitchell, n.d.

HOMOGENEITY OF VARIANCE ASSUMPTION (CONT.)

- Problem: You cannot determine which of these two values is responsible for a *t* statistic that falls in the critical region.
- There is no certainty that rejecting the null hypothesis is correct when obtaining an extreme value for *t*.
- If homogeneity of variance is violated, the following figure presents an alternative procedure for computing the *t* statistic that does not involve pooling the two sample variances



Source: Mitchell, n.d.

DMCC DIALOG MARKETING COMPETENCE CENTER

Two-Sample *t* Test for Homogeneous Variances

- Begin by calculating the mean and variance for each of your two samples.
- Then determine pooled variance S_p^2:

$$S_p^2 = \frac{\sum_{i=1}^{N_1}(y_i - \bar{y}_i)^2 + \sum_{j=1}^{N_2}(y_j - \bar{y}_j)^2}{(N_1 - 1) + (N_2 - 1)} = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{(N_1 + N_2 - 2)}$$

- Determine the test statistic *t*:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_p^2}{N_1} + \frac{S_p^2}{N_2}}} \qquad df = N_1 + N_2 - 2$$

Source: Ohio University, n.d.

4. Selecting Statistical Tests

Parametric Tests

DMCC Dialog Marketing Competence Center

F TEST

- ***F* test:** "A statistical test of the equality of the variances of two populations." (Malhotra, 2007, p. 481)

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

- ***F* statistic:** "The *F* statistic is computed as the ration of two sample variances." (Malhotra, 2007, p. 481)

$$F_{(n_1-1)(n_2-1)} = \frac{s_1^2}{s_2^2}$$

n_1 = size of sample 1

n_2 = size of sample 2

$n_1 - 1$ = degrees of freedom for sample 1

$n_2 - 1$ = degrees of freedom for sample 2

s_1^2 = sample variance for sample 1

s_2^2 = sample variance for sample 2

F TEST (CONT.)

- ***F* distribution:** "A frequency distribution that depends upon two sets of degrees of freedom – the degrees of freedom in the numerator and the degrees of freedom in the denominator." (Malhotra, 2007, p. 481)
- The critical values of *F* for various degrees of freedom for the numerator and denominator are given in statistical tables.
- If the probability of *F* is greater than the significance level α, H_0 is not rejected, and *t* based on the pooled variance estimate can be used.
- If the probability of *F* is less than or equal to α, H_0 is rejected and *t* based on a seperate variance estimate is used.

Source: Malhotra, 2007, p. 481.

F DISTRIBUTION

$\chi^2_{v_1}$ and $\chi^2_{v_2}$ are two independent χ^2-distributed random variables with v_1 and v_2 degrees of freedom.

The relation

$$F = \frac{\chi_1^2 / v_1}{\chi_2^2 / v_2}$$

is *F*-distributed with v_1 and v_2 degrees of freedom.

Source: Kosfeld, 2012b.

4. SELECTING STATISTICAL TESTS

PARAMETRIC TESTS

4.1	*T* TEST *F*
4.2	TEST CHI^2
4.3	TEST

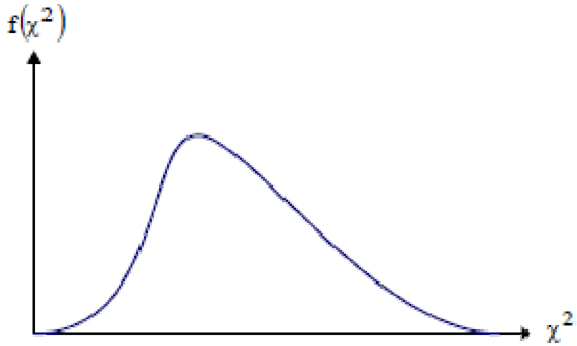
CHI² TEST

- Usually used to compare multiple groups where the input and the output variable are binary.
- "[…] used to test the statistical significance of the observed association in a cross-tabulation. It assists us in determining whether a systematic association exists between the two variables."
  (Malhotra, 2007, p. 474)

$Z_1, Z_2, \dots Z_n$, $\quad Z_i = \frac{X_i - \mu}{\sigma}$ are independent, standard normally-distributed random variables.

The sum of squares

$$(*)\ \chi^2 = Z_1^2 + Z_2^2 + \dots + Z_n^2 = \sum_{i=1}^{n} Z_i^2$$



Source: Malhotra, 2007, p. 474 f.; Neideen & Brasel, 2007, p. 94; Kosfeld, 2012b.

CHI² TEST (CONT.)

The sum of squares has a χ^2 distribution with *n* degrees of freedom. The sum of squares can also be presented as followed:

$$\chi^2 = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2}$$

Is μ estimated by the sample mean score → sum of squares:

$$\chi^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2}$$

χ^2 with *n-1* degrees of freedom. Generally a sum of squares of the form *(*)* is χ^2-distributed with n-k degrees of freedom. But only if there are *k* parameters of the sample to estimate.

Source: Kosfeld, 2012b.

EXAMPLE: CHI² TEST – GENERIC DATA

- 2×2 table of data
- Comparing two groups (one receiving a treatment, the other not)
- Comparing two outcomes of cure vs. continued disease

	Outcome		
Treatment	+	–	**Total**
Observed data			
+	45	20	65
–	5	30	35
total	50	50	100
Expected Data			
+	32.5	32.5	65
–	17.5	17.5	35
total	50	50	100

Source: Neideen & Brasel, 2007, p. 94 f.

EXAMPLE: CHI² TEST – GENERIC DATA (CONT.)

- Data table was constructed by taking the totals of the rows, multiplying them by the column totals, and then dividing by the grand total of subjects.
- Critical value is determined.
- The degrees of freedom for a chi-square table are the number of rows minus one times the number of columns minus one.
- For a 2×2 table, the degrees of freedom is always 1.

Example: calculated chi-squared statistic is 27.5, and the critical value is 3.841 → reject the H_0; the treatment positively affects outcome.

Source: Neideen & Brasel, 2007, p. 94 f.

4. Selecting Statistical Tests

Nonparametric Tests

4.4 Kolmogorov-Smirnov Test

KOLMOGOROV-SMIRNOV TEST

- A test of hypothesis that the sampled population follows some specified distribution.
- **K-S test** compares a continuous c.d.f.* *F(x)* to an empirical c.d.f $S_N(x)$, of the sample of N observations.
- If the sample from the random number generator is $R_1, R_2, \ldots, R_N$ → empirical c.d.f. $S_N(x)$ is given by:

$$S_N(x) = \frac{\text{number of } R_1,\ R_2,\ \cdots,\ R_N \text{ which are} \leq x}{N}$$

*c.d.f. = cumulative distribution function

Source: Aly, n.d.

KOLMOGOROV-SMIRNOV TEST (CONT.)

Step 1	Rank the data from the smallest to the largest: $R_1 \leq R_2 \leq \ldots \leq R_N$
Step 2	Compute: $D^+ = \max_{1 \leq i \leq N} \left\{ \frac{i}{N} - R_i \right\}$ $D^- = \max_{1 \leq i \leq N} \left\{ R_i - \frac{i-1}{N} \right\}$
Step 3	Compute: $D = max\ (D^+, D^-)$
Step 4	Locate the critical value in the “K-S Critical Values Table“* for the specified significance level α and the given sample size (N).
Step 5	If $D \leq D_\alpha$ → Accept “No difference between $S_N(x)$ and $F(x)$” If $D > D_\alpha$ → Reject “No difference between $S_N(x)$ and $F(x)$”

*Note: Critical Values Tables are found in many books.

Source: Aly, n.d.

Kolmogorov-Smirnov Critical Values Table

Degrees of Freedom (N)	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29
35	0.21	0.23	0.27
Over 35	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

Source: Massey, 1951, p. 70.

K-S Test Example

Five numbers: 0.44, 0.81, 0.14, 0.05, 0.93 were generated and it is required to test for uniformity using the K-S Test with the level of significance α = 0.05.

Solution:

R_i	0.05	0.14	0.44	0.81	0.93
i/N	0.2	0.4	0.6	0.8	1
$(i/N)-R_i$	0.15	0.26	0.16	---	0.07
$R_i-[(i-1)/n]$	0.05	---	0.04	0.21	0.13

Source: Aly, n.d.

K-S Test Example (cont.)

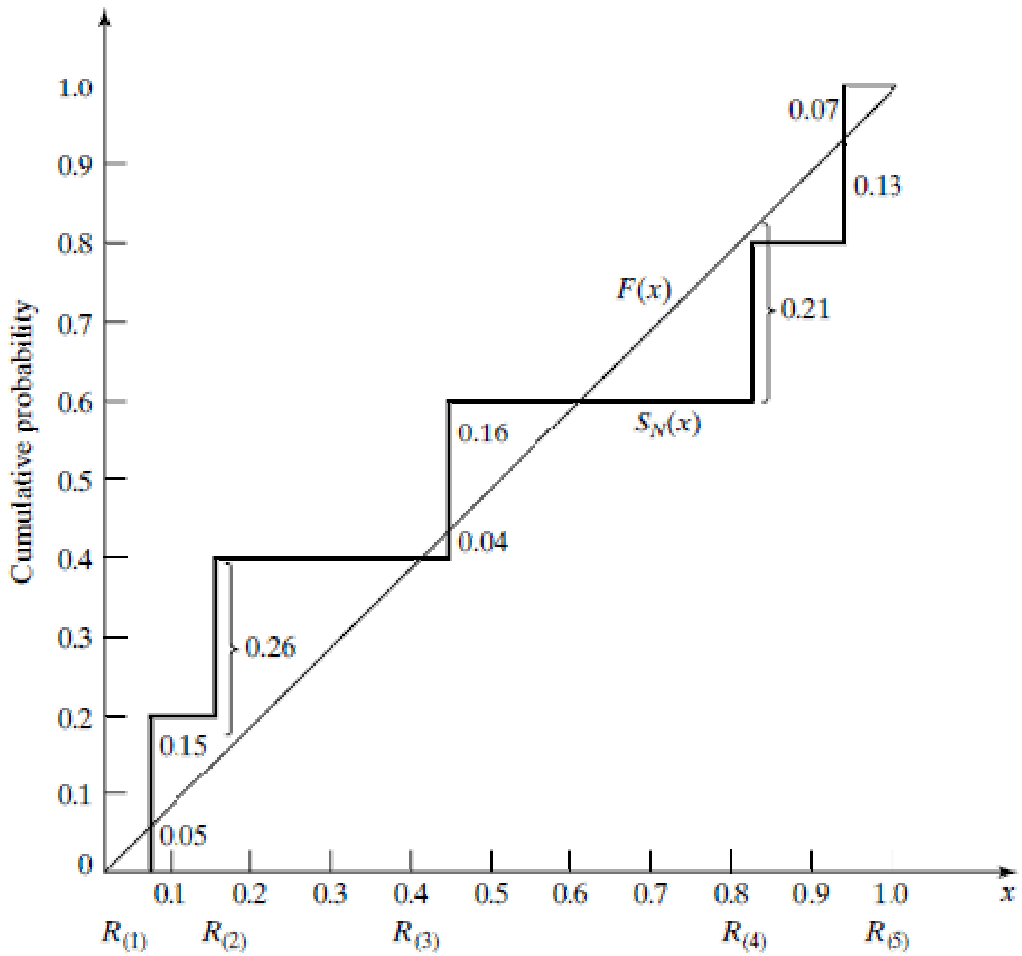
- $D = max(D^+, D^-) = max(0.26, 0.21) = 0.26$
- K-S Critical Values Table:

  $\alpha = 0.05$ and $N = 5$

  → Critical value $D_\alpha = 0.565$
- Since $D < D_\alpha$, no difference has been detected between the true distribution of $\{R_1, R_2, \ldots, R_N\}$ and the uniform distribution.

Source: Aly, n.d.

K-S TEST EXAMPLE (CONT.)



Source: Aly, n.d.

DMCC DIALOG MARKETING COMPETENCE CENTER

4. Selecting Statistical Tests

Nonparametric tests

4.4 Kolmogorov-Smirnov Test

4.5 Runs Test

Runs Test

- "The **runs test** is a test of randomness for the dichotomous variables. This test is conducted by determining whether the order or sequence in which observations are obtained is random."

  (Malhotra, 2007, p. 486)

- "The **runs test** examines the number of 'runs' of each of two possible characteristics that sample elements may have. A run is a sequence of identical occurrences of elements (symbols or numbers) preceded and followed by different occurrences of elements or by no element at all."

  (Sharma, 2008, p. 403)

RUNS TEST – EXAMPLE 1

In a bar, the bar keeper wishes to see whether the guests order their drinks (beer and cocktails) randomly. He observes the first 25 orders, with the following sequences of beer (B) and cocktails (C).

C	C	C	B	B	C	C	C	C	B	C	B	B	B	C	C	C	C	B	B	C	C	C	B	B

Test for randomness at $\alpha = 0.05$.

Source: Based on Bluman, 2007, p. 702.

DMCC DIALOG MARKETING COMPETENCE CENTER

RUNS TEST – EXAMPLE 1 (CONT.)

Step 1: State the hypotheses and identify the claim

- H_0: The guests order their drinks randomly, according to beverage (claim).
- H_1: The null hypothesis is not true.

Step 2: Find the number of runs

There are 15 cocktails and 10 beers.

Run	Beverage
1	CCC
2	BB
3	CCCC
4	B
5	C
6	BBB
7	CCCC
8	BB
9	CCC
10	BB

Source: Based on Bluman, 2007, p. 702.

RUNS TEST – EXAMPLE 1 (CONT.)

Step 3: Find the critical value

Find the number of runs for n_1 = 15, n_2 = 10, and α = 0.05. The values* are 7 and 18.

Step 4: Make the decision

Compare the critical values with the number of runs. Since the number of runs is 10, which is between 7 and 18, do not reject the null hypothesis.

Step 5: Summarize the results

There is not enough evidence to reject the hypothesis that the guest order their drinks randomly, according to beverage.

*Note: Critical Values Tables are found in many books.

Source: Based on Bluman, 2007, p. 702.

Runs Test – Example 2

- Tossing a coin 20 times produces the following sequence of heads (H) and tails (T):

HHH	TT	HH	TTT	HHH	TT	HHH	T
1	2	3	4	5	6	7	8

- First run of HHH is considered as run 1, second run of TT as run 2 etc.
- 8 runs = r = 8

Source: Sharma, 2008, p. 403.

Runs Test – Example 2 (cont.)

Small Sample

- n_1 = number of elements of one kind
- n_2 = number of elements of second kind
- $n = n_1 + n_2$ = total sample size
- $n_1 = 12$ heads
- $n_2 = 8$ tails
- One kind of elements is denoted by plus (+) sign, second kind of elements is denoted by minus (-) sign.
- + or – provides the direction of change from an established pattern.
- Is n_1 or n_2 less than 20, test is carried out by comparing the deserved number of runs R to critical values* of runs for given values of n_1 and n_2.

*Note: Critical Values Tables are found in many books.

Source: Sharma, 2008, p. 403 f.

Runs Test – Example 2 (cont.)

Small Sample

- H_0: Observations in the sample are randomly generated

  H_1: Observations in the samples are not randomly generated
- Can be tested that the occurrences of plus (+) signs and minus (-) signs are random by comparing *r* value with its critical value at a particular level of significance.
- *Decision Rule:*

  If $R \leq C_1$ or $R \geq C_2$ → Reject H_0

  Otherwise accept H_0
- C_1 and C_2 are critical values obtained from standard table with total tail probability $P(R \leq C_1) + P(R \geq C_2) = \alpha$.

Source: Sharma, 2008, p. 403 f.

DMCC DIALOG MARKETING COMPETENCE CENTER

RUNS TEST – EXAMPLE 2 (CONT.)

LARGE SAMPLE

- n_1 or $n_2 > 20 \rightarrow$ sampling distribution of R statistic (i.e. run) can be closely approximated by normal distribution.
- Mean and standard deviation of the number of runs for normal distribution are given by:

$$\text{Mean, } \mu_R = \frac{2n_1n_2}{n_1 + n_2} + 1$$

$$\text{Standard deviation, } \sigma_R = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

- Standard normal test statistic is given by:

$$Z = \frac{R - \mu_R}{\sigma_R} = \frac{r - \mu_R}{\sigma_R}$$

The critical z value is obtained in the usual manner at a specified level of significance.

Source: Sharma, 2008, p. 404.

4. Selecting Statistical Tests

Nonparametric tests

WILCOXON RANK SUM TEST

- Nonparametric comparison of two groups.
- Considers differences in magnitudes by using ranks.
- Used for independent samples.
- Used to compare distributions.
- Parametric equivalents: z and t test.
- Samples must be collected from approx. normally distributed population (unimodal).

Source: Bluman, 2007, p. 671.

WILCOXON RANK SUM TEST (CONT.)

- Values of the data for both samples are combined an than ranked.
- H_0 = true (no differences in population distribution)
  - Values in each sample should be ranked approx. the same.
  - Ranks are summed for each sample, sum should be equal → H_0 not-rejected.
- H_0 = rejected
  - Large difference in the sum of the ranks → distributions are not identical.

Source: Bluman, 2007, p. 671.

Formula for the Wilcoxon Rank Sum Test
(Independent Samples)

$$z = \frac{R - \mu_R}{\sigma_R}$$

where

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}$$

R = sum of ranks for smaller sample size (n_1)
n_1 = smaller of sample sizes
n_2 = larger of sample sizes
$n_1 \geq 10$ and $n_2 \geq 10$

Note that if both samples are the same size, either size can be used as *n_1*.

Source: Bluman, 2007, p. 672.

WILCOXON RANK SUM TEST – PROCEDURE TABLE

Step 1 State the hypotheses and identify the claim.

Step 2 Find the critical value(s).

Step 3 Compute the test value.

a. Combine the data from the two samples, arrange the combined data in order, and rank each value.

b. Sum the ranks of the group with the smaller sample size. (*Note:* If both groups have the same sample size, either one can be used.)

c. Use these formulas to find the test value.

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}$$

$$z = \frac{R - \mu_R}{\sigma_R}$$

where R is the sum of the ranks of the data in the smaller sample and n_1 and n_2 are each greater than or equal to 10.

Step 4 Make the decision.

Step 5 Summarize the results.

Source: Bluman, 2007, p. 673

Critical Values of the Wilcoxon Rank Sum Test

Two-Tailed Testing

- n = number of scores in the group with the smallest sum of ranks
- m = number of scores in the other group

n	α	m: 3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05			6	7	7	8	8	9	9	10	10	11	11	12	12	13	13	14
	.01	--	--	--	--	--	--	6	6	6	7	7	7	8	8	8	8	9	9
4	.05	--	10	11	12	13	14	14	15	16	17	18	19	20	21	21	22	23	24
	.01	--	--	--	10	10	11	11	12	12	13	13	14	15	15	16	16	17	18
5	.05	15	16	17	18	20	21	22	23	24	26	27	28	29	30	32	33	34	35
	.01	--	--	15	16	16	17	18	19	20	21	22	22	23	24	25	26	27	28
6	.05	22	23	24	26	27	29	31	32	34	35	37	38	40	42	43	45	46	48
	.01	--	21	22	23	24	25	26	27	28	30	31	32	33	34	36	37	38	39
7	.05	29	31	33	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62
	.01	--	28	29	31	32	34	35	37	38	40	41	43	44	46	47	49	50	52
8	.05	38	40	42	44	46	49	51	53	55	58	60	62	65	67	70	72	74	77
	.01	--	37	38	40	42	43	45	47	49	51	53	54	56	58	60	62	64	66
9	.05	47	49	52	55	57	60	62	65	68	71	73	76	79	82	84	87	90	93
	.01	45	46	48	50	52	54	56	58	61	63	65	67	69	72	74	76	78	81
10	.05	58	60	63	66	69	72	75	78	81	84	88	91	94	97	100	103	107	110
	.01	55	57	59	61	64	66	68	71	73	76	79	81	84	86	89	92	94	97
11	.05	69	72	75	79	82	85	89	92	96	99	103	106	110	113	117	121	124	128
	.01	66	68	71	73	76	79	82	84	87	90	93	96	99	102	105	108	111	114
12	.05	82	85	89	92	96	100	104	107	111	115	119	123	127	131	135	139	143	147
	.01	79	81	84	87	90	93	96	99	102	105	109	112	115	119	122	125	129	132
13	.05	95	99	103	107	111	115	119	124	128	132	136	141	145	150	154	158	163	167
	.01	92	94	98	101	104	108	111	115	118	122	125	129	133	136	140	144	147	151
14	.05	110	114	118	122	127	131	136	141	145	150	155	160	164	169	172	179	183	188
	.01	106	109	112	116	120	123	127	131	135	139	143	147	151	155	159	163	168	172
15	.05	125	130	134	139	144	149	154	159	164	169	174	179	184	190	195	200	205	210
	.01	122	125	128	132	136	140	144	149	153	157	162	166	171	175	180	184	189	193
16	.05	142	147	151	157	162	167	173	178	183	189	195	200	206	211	217	222	228	234
	.01	138	141	145	149	154	158	163	167	172	177	181	186	191	196	201	206	210	215
17	.05	159	164	170	175	181	187	192	198	204	210	216	220	228	234	240	246	252	258
	.01	155	159	163	168	172	177	182	187	192	197	202	207	213	218	223	228	234	239
18	.05	178	183	189	195	201	207	213	219	226	232	238	245	251	257	264	270	277	283
	.01	173	177	182	187	192	197	202	208	213	218	224	229	235	241	246	252	258	263
19	.05	197	203	209	215	222	228	235	242	248	255	262	268	275	282	289	296	303	309
	.01	193	197	202	207	212	218	223	229	235	241	246	253	259	264	271	277	283	289
20	.05	218	224	230	237	244	251	258	265	272	279	286	293	300	308	315	322	329	337
	.01	213	218	223	228	234	240	246	252	258	264	270	277	283	289	296	302	309	315

Source: Gdańsk University of Technology, n.d.

Critical Values of the Wilcoxon Rank Sum Test

One-Tailed Testing

- n = number of scores in the group with the smallest sum of ranks
- m = number of scores in the other group

n	α	m																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	6	6	7	8	8	9	10	10	11	11	12	13	13	14	15	15	16	17
	.01	--	--	--	--	6	6	7	7	7	8	8	8	9	9	10	10	10	11
4	.05	10	11	12	13	14	15	16	17	18	19	20	21	22	24	25	26	27	28
	.01	--	--	10	11	11	12	13	13	14	15	15	16	17	17	18	19	19	20
5	.05	16	17	19	20	21	23	24	26	27	28	30	31	33	34	35	37	38	40
	.01	--	--	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
6	.05	23	24	26	28	29	31	33	35	37	38	40	42	44	46	47	49	51	53
	.01	--	22	23	24	25	27	28	29	30	32	33	34	36	37	39	40	41	43
7	.05	30	32	34	36	39	41	43	45	47	49	52	54	56	58	61	63	65	67
	.01	28	29	31	32	34	35	37	39	40	42	44	45	47	49	51	52	54	56
8	.05	39	41	44	46	49	51	54	56	59	62	64	67	69	72	75	77	80	83
	.01	36	38	40	42	43	45	47	49	51	53	56	58	60	62	64	66	68	70
9	.05	49	51	54	57	60	63	66	69	72	75	78	81	84	87	90	93	96	99
	.01	46	48	50	52	54	56	59	61	63	66	68	71	73	76	78	81	83	85
10	.05	59	62	66	69	72	75	79	82	86	89	92	96	99	103	106	110	113	117
	.01	56	58	61	63	66	68	71	74	77	79	82	85	88	91	93	96	99	102
11	.05	71	74	78	82	85	89	93	97	100	104	108	112	116	120	123	127	131	135
	.01	67	70	73	75	78	81	84	88	91	94	97	100	103	107	110	113	116	119
12	.05	83	87	91	95	99	104	108	112	116	120	125	129	133	138	142	146	150	155
	.01	80	83	86	89	92	95	99	102	106	109	113	116	120	124	127	131	134	138
13	.05	97	101	106	110	115	119	124	128	133	138	142	147	152	156	161	166	171	175
	.01	93	96	100	103	107	111	114	118	122	126	130	134	138	142	146	150	154	158
14	.05	112	116	121	126	131	136	141	146	151	156	161	166	171	176	182	187	192	197
	.01	107	111	115	118	122	127	131	135	139	143	148	152	156	161	165	170	174	178
15	.05	127	132	138	143	148	153	159	164	170	175	181	186	192	197	203	208	214	220
	.01	123	127	131	135	139	144	148	153	157	162	167	171	176	181	186	190	195	200
16	.05	144	150	155	161	166	172	178	184	190	196	201	207	213	219	225	231	237	243
	.01	139	143	148	152	157	162	167	172	177	182	187	192	197	202	207	212	218	223
17	.05	162	168	173	179	186	192	198	204	210	217	223	230	236	242	249	255	262	268
	.01	157	161	166	171	176	181	186	191	197	202	208	213	219	224	230	235	241	246
18	.05	180	187	193	199	206	212	219	226	232	239	246	253	259	266	273	280	287	294
	.01	175	180	185	190	195	201	207	212	218	224	230	236	241	247	253	259	265	271
19	.05	200	207	213	220	227	234	241	248	255	262	270	277	284	291	299	306	313	320
	.01	194	199	205	210	216	222	228	234	240	246	253	259	265	272	278	284	291	297
20	.05	221	228	235	242	249	257	264	272	279	287	294	302	310	317	325	333	340	348
	.01	215	220	226	232	238	244	250	257	263	270	277	283	290	297	303	310	317	324

Source: Gdańsk University of Technology, n.d.

Wilcoxon Rank Sum Test

Example 1

Two independent samples of men and women are selected, and the time in minutes it takes each person to order the next beer is recorded, as shown in the table.

Men	15	18	16	17	13	22	24	17	19	21	26	28	Mean = 19,67
Women	14	9	16	19	10	12	11	8	15	18	25		Mean = 14, 27

At α = 0.05, is there a difference in the times it takes men/women to order the next beer?

Source: Based on Bluman, 2007, p. 672 ff.

WILCOXON RANK SUM TEST

SOLUTION – EXAMPLE 1

Step 1 State the hypotheses and identify the claim.

H_0: There is no difference in the times it takes men/women to order the next beer.

H_1: There is a difference in the times it takes men/women to order the next beer (claim).

Step 2 Find the critical value. Since $\alpha = 0.05$ and this test is a two-tailed test, use the z values* of +1.96 and -1.96.*(Note: Critical Values Tables are found in many books)

Step 3 Compute the test value.

a. Combine the data from the two samples, arrange the combined data in order, and rank each value. Be sure to indicate the group.

Time	8	9	10	11	12	13	14	15	15	16	16	17
Group	W	W	W	W	W	M	W	M	W	M	W	M
Rank	1	2	3	4	5	6	7	8.5	8.5	10.5	10.5	12.5
Time	17	18	18	19	19	21	22	24	25	26	28	
Group	M	W	M	M	W	M	M	M	W	M	M	
Rank	12.5	14.5	14.5	16.5	16.5	18	19	20	21	22	23	

.

Source: Based on Bluman, 2007, p. 672 ff.

WILCOXON RANK SUM TEST

SOLUTION – EXAMPLE 1 (CONT.)

Step 3 *b.* Sum the ranks of the groups with the smaller size (if both groups have the same size, either one can be used). Here: sample size for the women is smaller.
R = 1 + 2 + 3 + 4 + 5 + 7 + 8.5 + 10.5 + 14.5 + 16.5 + 21 = 93

c. Substitute in the formulas to find the test value.

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{(11)(11 + 12 + 1)}{2} = 132$$

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}} = \sqrt{\frac{(11)(12)(11 + 12 + 1)}{12}}$$

$$= \sqrt{264} = 16.2$$

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{93 - 132}{16.2} = -2.41$$

Step 4 Make the decision. The decision is to reject $H_{0,}$ since -2.41 < -1.96.

Step 5 Summarize the results. There is enough evidence to support the claim that there is a difference in the times it takes men/women to order the next beer.

Source: Based on Bluman, 2007, p. 672 ff.

WILCOXON RANK SUM TEST

EXAMPLE 2

Main Idea: If two groups come from the same distribution, but you've just randomly assigned labels to them, values in the two different groups should have values somewhat equally distributed between the two.

Group A: $X_1,\ldots, X_{n_1} \sim F_A$

Group B: $Y_1,\ldots, Y_{n_2} \sim F_B$

H_0: $F_A = F_B$

Group A	1.3	3.4		$n_1 = 2$
Group B	4.9	10.3	3.3	$n_2 = 3$

Order all observations in the combined sample and assign ranks:

Order	1.3	3.3	3.4	4.9	10.3
Assign ranks	1	2	3	4	55

Source: Holmes, 2004.

WILCOXON RANK SUM TEST

EXAMPLE 2 (CONT.)

Test statistic R_1 = sum of ranks attached to Group A = 1 + 3 = 4

Under H_0, each 2-subset of the ranks *{1,2,3,4,5}* is equally likely to occur as the ranks of X_1, X_2.

Possible ranks for X_1, X_2:

sum =	R_1
{1,2}	3
{1,3}	4
{1,4}	5
{1,5}	6
{2,3}	5
{2,4}	6
{2,5}	7
{4,5}	9
{3,4}	7
{3,5}	8

Source: Holmes, 2004.

Wilcoxon Rank Sum Test

Example 2 (cont.)

Hence the distribution of R_1 under H_0 is given by:

$r =$	3	4	5	6	7	8	9
$P_{H_0}(R_1 = r)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{10}$

$R_1 = 4$, the one-sided P-value

$P = P_{H_0}$ (seeing a value as small $(R_1 \leq 4) = P$ or smaller than observed) $= \frac{1}{5}$

Source: Holmes, 2004.

5. Common Statistical Errors

MEASUREMENT ACCURACY

- **Measurement error:**

  "The variation in the information sought by the researcher and the information generated by the measurement process employed."
  (Malhotra, 2007, p. 283)

- Variety of factors can cause **measurement errors**, which result in the measurement/observed score being different from the true score of the characteristic being measured.
- **True score model** provides a framework for understanding the accuracy of measurement:

  $X_O = X_T + X_S + X_R$

  X_O = the observed score or measurement
  X_T = the true score of the characteristic X_S
  = systematic error
  X_R = random error

Source: Malhotra, 2007, p. 283 f.

Measurement Accuracy (cont.)

- **Systematic error:**

  "Affects the measurement in a constant way and represents stable factors that affect the observed score in the same way each time the measurement is made."

  (Malhotra, 2007, p. 284)

- **Random error:**

  "Measurement error that arises from random changes or differences in respondents or measurement situations."

  (Malhotra, 2007, p. 284)

POTENTIAL SOURCES OF ERROR IN MEASUREMENT

1. Other relatively stable characteristics of the individual that influence the test score (e.g., intelligence, social desirability, education).
2. Short-term or transient personal factors (e.g., health, emotions, fatigue).
3. Situational factors (e.g., presence of other people, noise, distractions).
4. Sampling of items include in the scale: addition, deletion, or changes in the scale items.
5. Lack of clarity of the scale, including the instructions or the items themselves.
6. Mechanical factors (e.g., poor printing, overcrowding of items in questionnaire, poor design).
7. Administration of the scale (e.g., differences among interviewers).
8. Analysis factors (e.g., differences in scoring and statistical analysis).

Source: Malhotra, 2007, p. 284.

DESIGN ERRORS

- **Sample size for human subjects:**

Many studies are too small to detect even large effects.

Guide to sample size

Expected difference (p1-p2)	Total sample size required*
5%	1450-3200
10%	440-820
20%	140-210
30%	80-100
40%	50-60

* 5% significance level, 80% power. Smaller numbers may be justified for rare outcomes (p1 <.1)

Source: Brady, n.d.

DESIGN ERRORS (CONT.)

- **Method comparison studies:**

  If different methods are evaluated by different observers then the method differences are confounded with observer differences. The study must be repeated with each observer using all methods.

Source: Brady, n.d.

DESIGN ERRORS (CONT.)

- **Bias:**
  - Randomization is the best way of avoiding bias but it is not always possible or appropriate.
  - Appreciation and measures taken to reduce bias through study design.
  - Collection of data, definition and assessment of outcome and, method of randomization should be clearly described.
  - Appropriate analytic methods such as multiple regression should be used to adjust for differences between groups in observational studies.
  - Authors should discuss likely biases and potential impact on their results.

Source: Brady, n.d.

Analysis Error

- Failure to use a test for trend on ordered categories (e.g. age group).
- Dichotomizing continuous variables in the analysis (acceptable for descriptive purposes).
- Using methods for independent samples on paired or repeated measures data.
- Using parametric methods without verifying for normal distribution.
- Over using hypothesis tests (*p*-values) in preference to confidence intervals.
- One-tailed tests are very rarely appropriate.
- Obscure statistical tests should be justified and referenced.

Source: Brady, n.d.

ANALYSIS ERROR (CONT.)

- Comparing *p*-values between subgroups instead of carrying out tests of interaction is incorrect.
- Correlating time series: Any two variables that consistently rise, fall or remain constant over time will be correlated. 'Detrended' series should be compared instead.
- **Multiple testing:**
  Conclusions should only be drawn from appropriate analyses of a small number of clear, pre-defined hypotheses. Results from post-hoc subgroup/risk-factor analyses should be treated as speculative. If many such tests have been carried out adjustment for multiple testing should be considered. Comparing groups at multiple time points should be avoided – a summary statistics approach or more complex statistical methods should be used instead.

Source: Brady, n.d.

Common Statistical Mistakes

Mistake 1:

- Failing to investigate data for data entry or recording errors.
- Failing to graph data and calculate basic descriptive statistics before analyzing data.

Mistake 2:

- Using the wrong statistical procedure in analyzing your data.
- Includes failing to check that necessary assumptions are met.

Mistake 3:

- Failing to design your study so that it has high enough **power** to call meaningful differences "significantly different".
- Includes concluding that the H_0 is true. Should be "not enough evidence to say the H_0 is false."
- Confidence interval tells you what the population value might be.

Source: Simon, 1999.

COMMON STATISTICAL MISTAKES (CONT.)

Mistake 4:

- Failing to report a confidence interval as well as the *P*-value.
- *P*-value tells you if statistically significant.

Mistake 5:

- "Fishing" for significant results. That is, performing several hypothesis tests on a data set, and reporting only those results that are significant.
- If α = P(Type I) = 0.05, and we perform 20 tests on the same data set, we can expect to make 1 Type I error. ($0.05 \times 20 = 1$).

Mistake 6:

- Overstating the results of an observational study.
  - That is, suggesting that one variable "caused" the differences in the other variable.
  - As opposed to correctly saying that the two variables are "**associated**" or "**correlated**."
- Don't forget that a significant result may be "spurious."

Source: Simon, 1999.

COMMON STATISTICAL MISTAKES (CONT.)

Mistake 7:

- Using a non-random or unrepresentative sample.
- Includes extending the results of an unrepresentative sample to the population.

Mistake 8:

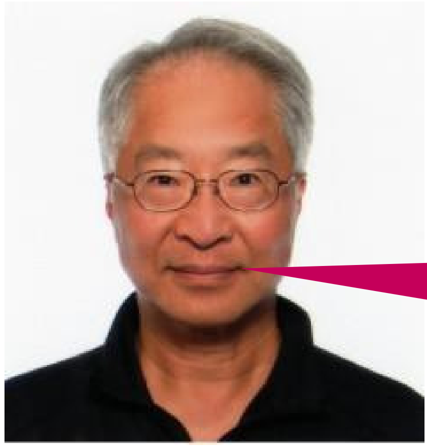
- Failing to use all of the basic principles of experiments, including randomization, blinding, and controlling.

Source: Simon, 1999.

6.

Excursus: Nothing can be proved Empirically

PRACTICE VS. ACADEMIC RESEARCH



"Much academic research on information technology, systems, and management has been branded by practitioners in business as unusable, irrelevant, and unreadable."

(Ho, 2000, p. 6)

"The more I read in the Journal of Marketing, or tried to read in the Journal of Marketing Research, the more I realized that what was meaningful to the marketing academician in terms of tools or techniques to solve ma marketing problems, had little or no application in the marketplace."

(Maiken, 1979, p. 58)

Practice vs. Academic Research (cont.)

"Is academic research in business management disciplines readily related to workplace issues and practical management skills? And is it typically conveyed in terms familiar to practitioners. Generally, the answer to both would seem to be no."

(Wilkerson, 1999, p. 599)

"Throughout its 100-plus year history, one of the most recurring themes has been that there is a 'gap' or 'divide' between marketing academe and marketing practice. As evidence, critics point out (among other things) that marketing practitioners neither subscribe to nor read academic marketing journals."

(Hunt, 2002, p. 305)

DMCC DIALOG MARKETING COMPETENCE CENTER

SEVEN REASONS WHY PRACTITIONERS SHOULD IGNORE ACADEMIC RESEARCH

1. Customers
2. Structure
3. Causality
4. Reductionism
5. Precision
6. Generalizations
7. Replication

Source: November, 2004, p. 41 ff.

1. CUSTOMERS

- Practitioners are **NOT** the customers for academic research!
- They **DON'T** ask academic scientist to do the research and they **DON'T** pay academic scientist to do the research.
- There is **no reason** why practitioners should expect academics to produce something of value to them.
- Number of publications and type of journal are important to most academics/universities.
- Article with the title "Aspects of Chi Square Testing in Structural Equation Modeling" published in the *Journal of Marketing Research* would be regarded as much more valuable as an article with the title "How to Make Your Web Site Irresistible" published in *Marketing Magazine*.

Source: November, 2004, p. 41.

2. STRUCTURE

- Practitioners tend to use their own personal practice as a frame of reference.
- Principle of academic freedom of enquiry means that each academic decides for himself what to research.
- Publication matters – **NOT** whether the research will be useful for practitioners.
- Thousands of independent and capricious studies are produced which result in an arbitrary, chaotic and unpredictable collection of work that has no apparent structure.
- How are practitioners supposed to put any given paper into a general context of meaning when academic scientists have no such thing?

Source: November, 2004, p. 41 f.

3. CAUSALITY

- Academics make sometimes false or misleading statements about causality in their arguably misguided efforts in seeking relationships between variables in marketing systems.
- Academics should understand the difference between causality and association.
- *Example:* "The findings support our hypothesis that […] market orientation is an important determinant of profitability."
  - Unconditional causal statement; misleading.
  - Readers are given no indication of how important market orientation is, nor what other things are important, nor how it compares with other orientations.
  - No explanation how much effort should be put into obtaining a market orientation, nor the extend to which the cost of doing this will affect profitability.

Source: November, 2004, p. 42 f.

4. Reductionism

- There is a good chance that practitioners will not appreciate the dangers inherent in studying small parts of systems and then applying the knowledge gained to other parts or to the system as a whole.
- Academics should never generate generalizations from a single reductionist study.
- Conclusions are statistically valid only for the study itself, it does not extend to other people, countries, websites, time periods or products.
- Practitioner could misuse the study unless it contains a clear warning of its limitations.

Source: November, 2004, p. 43.

5. Precision

- Practitioners might be deluded into thinking that it is thick ice when in fact it is thin.
- Authors mislead their readers (and themselves) into thinking that their results are more meaningful than they really are.
- E.g., poor data can never be corrected by high statistical validity. Measurement systems are the weakest part of the researchers work and the weakness cannot be corrected after measurements have been made.
- It is perilously easy to create a false sense of precision through the application of statistical validity tests.

Source: November, 2004, p. 43 ff.

6. GENERALIZATIONS

- The few generalizations academics which produce largely only corroborate what practitioners already know.
- They do not help with decision detail and can become unquestioned articles of faith.
- Marketing generalizations – e.g., positive relationship between advertising spending and sales – do not help with the specifics of advertising decision.
- There are innumerable problems of detail to be handled, any of which can affect success.

Source: November, 2004, p. 45 f.

7. Replication

- The truth-value of academic research is highly questionable – often it is not even thin ice, it is water.
- Principle of replication has been acknowledged as an essential part of scientific study.
- Even after a study which satisfies the stringent selection criteria of the leading journals has been published, there are no reasons for accepting the findings until the work has been independently replicated a reasonable number of times without major anomaly.
- Unreplicated work is virtually meaningless and useless irrespective of its level of statistical significance.

Source: November, 2004, p. 46 f.

CONCLUSION

"[…]managers **should continue to avoid** reading academic work and avoid attending academic conferences because we are simply **not** yet **able to produce knowledge** that is **useful to practitioners** in anything other than very meagre quantities. It is very **difficult** for practitioners to find this work among the **very large quantity** that, quite rightly, **only serves academic ends**.

Until a work has been thoroughly replicated it should contain, not a section on managerial implications, but a section giving a **clear warning against the application of the work in any practical circumstance**. Although academics are often criticised for writing in an **abstruse style**, this is actually an **advantage** since it is **less likely** that **practitioners will read** the work." (November, 2004, p. 47)

THANK YOU VERY MUCH FOR YOUR ATTENTION!

LITERATURE

- **Aaker, D.A., Kumar, V. & Day, G.S. (2007).** Marketing Research (9th ed.). USA: Wiley.
- **Access Excellence (n.d.).** Writing Hypotheses: a student lesson. Accessed on 25th February 2013 via http://www.accessexcellence.org/LC/TL/filson/writhypo.php.
- **American Marketing Association (AMA) (2013a).** Definition of Behavioral Intention. Accessed on 27th February 2013 via http://www.marketingpower.com/_layouts/dictionary.aspx?dLetter=B.
- **American Marketing Association (AMA) (2013b).** Definition of Purchase Intention. Accessed on 27th February 2013 via http://www.marketingpower.com/_layouts/dictionary.aspx?dLetter=P.
- **American Marketing Association (AMA) (2013c).** Definition of Planned Buying vs. Unplanned Buying. Accessed on 27th February 2013 via http://www.marketingpower.com/_layouts/dictionary.aspx?dLetter=P.

LITERATURE

- **American Marketing Association (AMA) (2004).** Definition of Marketing Research. Accessed on 22nd February 2013 via http://www.marketingpower.com/aboutama/pages/definitionofmarketing.aspx.
- **Aly, A.I. (n.d.).** Test for Random Numbers. Accessed on 12th February 2013 via http://eee.guc.edu.eg/Courses/Networks/NETW707%20Modelling%20&%20Simulation/Lectures/Lecture%209%20-%20Tests%20for%20Random%20Numbers.pdf.
- **Bass, F.M. & Wind, J. (1995).** Introduction to Special Issue: Empirical generalizations in Marketing, Marketing Science, 14, p. 1-5.
- **Bluman, A.G. (2007).** Elementary Statistics (6th ed.). USA: McGraw-Hill.
- **Bortz, J. & Döring, N. (2006).** Forschungsmethoden und Evaluation (4th ed.). Heidelberg: Springer.

LITERATURE

- **Brady (n.d.).** Reviewer's quick guide to common statistical errors in scientific papers. Accessed on 14th February 2013 via http://cdn.elsevier.com/assets/pdf_file/0008/110996/reviewers_statistics.pdf.
- **Brett, M.; Penny, W. & Kiebel, S. (2003).** An Introduction to Random Field Theory. Accessed on 7th February 2013 via http://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf2/pdfs/Ch14.pdf.
- **Bronstein, I.N. & Semendjajew, K.A. (1983).** Taschenbuch der Mathematik. Frankfurt: Harri deutsch.
- **Bunge, M. (1967).** Scientific Research: The Search for Information. New York: Springer.
- **Centre for Innovation in Mathematics Teaching (n.d.).** Hypothesis Testing: One Sample Tests. Accessed on 29th January 2013 via http://www.cimt.plymouth.ac.uk/projects/mepres/alevel/fstats_ch3.pdf.
- **Decker, R. & Wagner, R. (2002).** Marketingforschung. München: Redline Wirtschaft.

LITERATURE

- **Eid, M., Gollwitzer, M. & Schmitt, M. (2010).** Statistik und Forschungs-methoden. Weinheim: Beltz.
- **Food and Agriculture Organization of the United Nations (FAO) (n.d.).** The Role of Marketing Research. Accessed on 16th January 2013 via http://www.fao.org/docrep/W3241E/w3241e02.htm#step%202:%20hypothesis%20generation.
- **Gdańsk University of Technology (n.d.).** Critical Values of the Wilcoxon Rank Sum Test. Accessed on 24th February via http://www.zie.pg.gda.pl/c/document_library/get_file?p_l_id=486577&folderId=486635&name=DLFE-7316.pdf.
- **Ho, J.K. (2000).** Bridging academic research and business practice with the new media. Information Resources Management Journal, 13 (2), p. 6-14.
- **Holmes, S. (2004, 12th July).** Non-Parametric tests. Accessed on 30th January 2013 via http://www-stat.stanford.edu/~susan/courses/s141/hononpara.pdf.

LITERATURE

- **Hunt, S.D., (2002).** Marketing as a profession: On closing stakeholder gaps. European Journal of Marketing, 36 (3), p. 305-312.
- **Kosfeld, R. (2012a).** Tabellen zu Prüfverteilungen. Accessed on 24th February 2013 via http://www.ivwl.uni-kassel.de/kosfeld/lehre/multivariate/skript/Verteilungstabellen.pdf.
- **Kosfeld, R. (2012b).** Exkurs:Chi-Quadrat- und F-Verteilung. Accessed on 7th February 2013 via http://www.ivwl.uni-kassel.de/kosfeld/lehre/multivariate/skript/Exkurs_Chi2-_und_F-Verteilung.pdf.
- **Kapanadze, S. (n.d.).** Examples of Good and Bad Hypotheses. Accessed on 25th February 2013 via http://de.scribd.com/doc/83770195/Examples-of-Good-and-Bad-Hypotheses.
- **Kalwani, M.U. & Silk, A.J. (1982).** On the Reliability and Predictive Validity of Purchase Intention Measures. Marketing Science, 1 (3), p. 243-286.

LITERATURE

- **Lilien, G.L., Kotler, P. & Moorthy, K.S. (1992).** Marketing Models. New Jersey: Prentice-Hall.
- **Lingayas Institute of Management and Technology (LIMAT) (n.d.).** Research Methodology: An Introduction. Accessed on 29th January 2013 via http://www.limat.org/data/research/Research%20Methodology.pdf.
- **Maiken, J.M. (1979).** What is the appropriate orientation for the marketing academician? In Ferrell et al. (Eds.), Conceptual and Theoretical Developments in Marketing (p. 58). Chicago: American Marketing Association.
- **Malhotra, N.K. (2007).** Marketing Research (5th ed.). New Jersey: Pearson.
- **Malhotra, N., Peterson, M. & Kleiser, S.B. (1999).** Marketing research: A state-of-the-art review and directions for the twenty-first century. Journal of the Academy of Marketing Science, 27 (Spring), p. 160-183.
- **Massey, F.J. (1951).** The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association, 46 (253), p.68-78.

LITERATURE

- **Mitchell, T. (n.d.).** Chapter 10: The t test for two independent samples. Accessed on 25th February 2013 via http://home.ubalt.edu/tmitch/631/PowerPoint_Lectures/chapter10/.
- **Neideen, T. & Brasel, K. (2007).** Understanding Statistical Tests. Journal of Surgical Education, 64(2), p. 93-96.
- **November, P. (2004).** Seven Reasons Why Marketing Practitioners Should Ignore Marketing Academic Research. Australasian Marketing Journal, 12 (2), p. 39-50.
- **Nowak, R. (2010).** Lecture 12: Multiple Testing. Accessed on 7th February 2013 via http://nowak.ece.wisc.edu/ece830/ece830_lecture12.pdf.
- **Ohio University (n.d.).** T-Test Statistics. Accessed on 25th February 2013 via http://www.ohio.edu/plantbio/staff/mccarthy/quantmet/lectures/ttest.pdf.
- **Palazzolo, D. & Roberts, D. (2010).** What is a Good Hypothesis? Accessed on 25th February 2013 via http://writing2.richmond.edu/writing/wweb/polisci/hypothesis.html.

LITERATURE

- **Popper, K.R. (1989).** Logik der Forschung. Tübingen: Mohr Siebeck.
- **Redman, L.V. & Mory, A.V.H. (1933).** The Romance of Research. Baltimore: The Williams & Wilkins Company.
- **Sekaran, U. & Bougie, R. (2010).** Research Methods for Business: A Skill Building Approach (5th ed.). UK: John Wiley & Sons.
- **Sage Publications (n.d.).** How Can I Create a Good Research Hypothesis? Accessed on 25th February via http://www.sagepub.com/upm-data/41397_32.pdf.
- **Simon, L.J. (1999, 12th June).** Common Statistical Mistakes. Accessed on 14th February 2013 via http://sites.stat.psu.edu/~lsimon/stat250/fa99/slides/mistakes/index.htm.
- **Solomon, M, Bamossy, G., Askegaard, S. & Hogg, M.K. (2006).** Consumer Behaviour (3rd ed.). England: Pearson.

LITERATURE

- **University of West England (2007).** Developing research and scientific hypotheses. Accessed on 25th February 2013 via http://ro.uwe.ac.uk/RenderPages/RenderLearningObject.aspx?Context=6&Area=1&Room=3&Constellation=54&LearningObject=332.
- **Wilkerson, J.M. (1999).** On research relevance, professors' "real world" experience, and management development: are we closing the gap? The Journal of Management Development, 18 (7), 598-608.
- **Williams, K.C. (1992).** Behavioural aspects of marketing. Oxford: Heinemann.