Bulletin
Volume 24/ No. 4
25/No. 1
March, June 2002

Editor's Foreword

First of all, I would like to apologize for the belated issue of this bulletin. Secondly, I would like to welcome Ms. Teresa Ng of City University of Hong Kong to join the Editorial Board as an Associate Editor.

In this issue, Prof. W.K. Li gives us an update on the work of the Society in the President's Forum as usual.

In February, a half-day workshop has been organized by the Professional Service Sub-committee. Dr. Chan Wai, the Chairman of the Sub-committee, writes a report on the workshop.

One of the speakers in the Workshop, Dr. Darius Chan of Chinese University of Hong Kong kindly writes an article on assessment tools, which is also the topic he discussed in the Workshop. We will invite other speakers to provide us with articles relating to their talks in the Workshop and they will appear in the coming issues.

Mr. Jonathan Ng of Census and Statistics Department writes an article on Sampling Design of the General Household Survey. Members should find it very useful and practical.

I write a book review on *Lady Tasting Tea* by David Salsburg, a retired statistician from Pfizer Pharmaceuticals. It is a book about the development of modern statistics. There are a lot of stories about famous statisticians. You will find out who introduced the student's t statistic and why it is named as student's t statistic.

We hope that you find it enjoyable to read the Bulletin. The Bulletin won't be success without your support.

P S Chan

Editor	:	Chan, Ping-shing, CUHK	Tel.	2609 7920	Fax.	2603 5188
Associate Editors	:	Chan, Wai, CUHK		2609 6241		2603 5019
		Chan, Wai-sum, HKU		2857 8318		2858 9041
		Ms Ng, Teresa, CityU		2788 8565		27888560
Secretary	:	Lam, John Hon-kwan, C&SD		2802 1267		2121 8296

CONTENTS

(Vol. 24/No.4 Vol. 25/No. 1, March, June 2002)

	Page
President's Forum *W.K. Li*	1
Report on Workshop on Social Survey and Research *Wai Chan*	2
Assessment Tools in Organizational Setting *Darius Chan*	4
Sampling Design of the General Household Survey *Jonathan Ng*	11
Book Review – The Lady Tasting Tea : How Statistics Revolutionized Science in the Twentieth Century By David Salsburg *P.S. Chan*	16
News	18

President's Forum

Professor W.K. Li

I apologize for this belated issue of the Bulletin because of events unforeseeable. Nevertheless, I hope you will find this issue interesting as it has collected two articles from a half-day workshop on social survey organized earlier by our Consultation Services Secretary, Dr. Chan Wai. This issue contains also a book review by Dr. Chan Ping-shing on an interesting book on the history of modern statistics, 'The Lady Tasting Tea' by David Salsburg.

I would like to take this opportunity to thank Dr. Mike So for his services to the Society as Membership Secretary for over three years and to welcome Dr. Li Leong Kwan who has come on board to take up Mike's job. The HKSS professional examination has a strong going, attracting some 70 examination associates, thanks to the effort of Mr. Fung Hing Wang and his team.

In the coming year we will be celebrating the 25th Anniversary of the Society. A half-day workshop on financial and economic statistics is scheduled tentatively in February, 2003.

I look forward to your continual support of the Society.

Report on Workshop on Social Survey and Research

Wai Chan
Consultation Services Secretary

A half-day workshop "Workshop on Social Survey and Research" was organized by the Society on February 23, 2002. Four speakers, Prof. Darius Chan from CUHK, Prof. Chan Ying Keung from CUHK, Prof. Tong Fung from HKU, and Mr. Alvin Li from the Census and Statistics Department, were invited to present in the Workshop. The workshop was very successful and it had attracted about 100 participants from different institutions. Some photos had been taken and presented here.

The speaker Prof. Darius Chan

The Speaker, Prof. Fung Wing Kam

HKSS President, Prof. Li Wai Keung presented a souvenir to Mr. Alvin Li

HKSS President, Prof. Li Wai Keung presented a souvenir to Prof. Chan Ying Keung

Bank for its generosity in providing us the free use of its lecture theatre.

The audience asked questions during the discussion period

Finally I would like to thank Hang Seng

Assessment Tools in Organizational Setting

Invited speech delivered in the Annual Workshop on Social Survey and Research, Hong Kong Statistical Society (Feb 23rd, 2002)

Prof. Darius Chan
The Chinese University of Hong Kong

Introduction

In a recent recruitment of ten senior administrative officers, the Government of Hong Kong S.A.R. spent hundreds of thousand of dollars to hire a consulting firm to adopt the "assessment center" approach to conduct the recruitment (Feb 20th, 2002, Ming Pao). While the concept of assessment center has been introduced in the personnel selection literature for quite some time, the practice is still rather uncommon for a lot of organizations in Hong Kong. The fact that the S.A.R. Government made the news because of using of this selection method serves as a good example. In the following, I will briefly describe the characteristics of three types of assessment, namely, selection interviews, psychometric tests, and assessment center. I will also mention some basic concepts in personnel assessment, and highlight the important factors that practitioners should pay attention to when deciding which assessment tool(s) to use. While these assessments are discussed in the context of personnel selection, they are also applicable to other kinds of personnel assessment, such as for promotion or career development.

I. Basic Concepts in Personnel Assessment

A. Reliability and Validity

There are many different types of personnel selection tools. The commonly used methods include selection interviews, psychometric tests (paper-and-pencil or computer-based testing), behavioral assessment, performance tests, assessment center, reference check, bio-data, etc. While some of these are more commonly used in actual practice (e.g., interviews) than others, we should be aware of the fact that each of these selection tools has its strengths and weaknesses.

In the context of personnel selection, two basic concepts are important for considering which assessment tool(s) to use, namely, reliability and validity. Reliability generally refers to how consistent and dependable an assessment tool is. Validity refers to the degree to which available evidence supports inferences made from scores on that

assessment tool (see Gatewood & Feild, 1997 for a detailed discussion). For instance, an assessment tool is considered to be valid if the assessment scores can accurately predict future job performance (as a criterion) of the selected candidates. In fact, validity is typically considered to be the most important characteristic of measures used in personnel selection.

When considering the validity of an assessment tool, there are a number of different strategies for obtaining the evidence to see if the inferences based on that assessment tool are accurate and can be supported. Here I will focus on *criterion-related* validity, which refers to the effectiveness of an assessment tool in predicting an individual's performance in some specified criterion (e.g., job performance as a criterion measure). An assessment tool is typically considered as valid if the relation between the assessment score and the criterion measure is found to be statistically significant (usually determined by a correlation coefficient). Two types of criterion-related validity are frequently examined in validity research. When the information is obtained on both a predictor (e.g., an assessment test score) and a criterion (e.g., job performance record) for a current group of employees, the validity score derived is known as *concurrent* validity. When the information of the assessment score is used instead to predict some specified criterion over time, the validity score derived is known as *predictive* validity. In the following discussion, I will predominantly use predictive validity when talking about the psychometric properties of various assessment tools.

B. When is Assessment Important - a Utility Perspective

Before we examine each of the three assessment tools, I would like to highlight the importance of choosing the "right" assessment using a utility perspective. Numerous researchers have proposed various ways to quantify the practical utility of using some more valid assessment tools over others. According to Brogden (1949) and Schmidt and associates (Schmidt et al., 1979; Schmidt & Hunter, 1998), the practical gains per person hired per year on the job after using a new selection method (denoted as ΔUtility/hire/year) is a three-way product:

$$\Delta\text{Utility/hire/year} = \Delta r_{xy} * SD_y * Mean\ Z_x$$

where Δr_{xy} is the difference between the validity of the new selection method and the old selection method;
SD_y refers to the standard deviation across individuals of the dollar value of job output;
$Mean\ Z_x$ refers to the average score on the employment procedure of those hired (in z-score form), as compared to the general applicant pool

For instance, if the predictive validity coefficient of the new selection method is .5 while the old method is .3, Δr_{xy} is then .2. As for SD_y, research in the U.S. suggests that it is typically found to be at minimum 40% of the mean salary. If we consider a job of average annual salary of $500,000, SD_y in this case becomes $200,000. Let us also assume that the selected applicants have a mean z score of +2 when compared to the applicant pool, then the practical gains per person hired per year using this new method is .2 * $200,000 * 2 = $80,000. Notice that this gain is calculated per person hired per year. This simple illustration provides the reader with some idea of how to quantify the practical gains brought along by an improved selection method. With the costs of the new selection method in mind, practitioners can then decide how to choose a selection method.

II. Three Types of Assessment Tools

As mentioned above, there are many different types of assessment tools. In this section, I will focus on three types, namely, selection interviews, psychometric tests, and assessment center, and briefly describe the characteristics of each of these methods.

A. The Selection Interview

The selection interview has long been considered as the most frequently used selection tool. Because of its widespread use, selection researchers have studied this tool for over 70 years. Despite its popularity, results from recent meta-analytic studies raise concern for its use. Specifically, selection interviews can be further classified into two broad categories, namely *structured* interviews and *unstructured* interviews. In structured interviews, questions are usually predetermined and interviewees' responses are scored in some predetermined fashion (e.g., based on some scoring keys). In other words, this type of interview is structured according to some *a priori* examination, e.g., through a job analysis, to identify important qualities that the interviewer would look for through the interview. The highly structured format also tries to minimize the potential biases of the interviewer. In unstructured interviews, the interviewer can ask any question that he/she wishes and the logic of the questions is up to the interviewer. This latter type of interview also makes up the majority of interviews being conducted in organizations. According to one recent meta-analysis (Schmidt & Hunter, 1998), the predictive validity of structured interviews for future job performance is .51, while that of the unstructured interviews is only .38.

Many of the problems of the selection interview have been identified in unstructured situations. Here I will just briefly mention some frequently mentioned biases in the interview literature and simple suggestions for minimizing the impact of such biases:

1. *Preconceptions and Stereotypes* – It refers to the interviewer's tendency to stereotype the applicant and allow personal bias to influence evaluations. Thus, the interviewer should be reminded of such a tendency and should make optimal use of the biographical data of the interviewee.
2. *Primacy effects* – It refers to the interviewer's tendency to make an evaluation of the applicant within the first few minutes of the interview. It is also known as the "first impression error". Thus, the interviewer should take into account all observed performances of the interviewee.
3. *Halo effects* – It refers to the interviewer's tendency to over-generalize the performance of the interviewee in one aspect to other aspects. Thus, the interviewer should judge the interviewee in all major aspects independently.
4. *Reciprocity bias* – It refers to the common phenomenon that performance of the interviewee is often affected by the interviewer. Thus, the interviewer should interpret a specific behavior in light of the specific context.
5. *Contrast effect* – It refers to the interviewer's tendency to allow the quality of the applicants who preceded the present applicant to influence the ratings of the present applicant. Thus, the interviewer should instead use all previous candidates as the anchor.
6. *Negative bias* – It refers to the interviewer's tendency to overemphasize negative behavior. Thus, the interviewer should balance negative and positive performances of the interviewee.

B. Psychometric Tests

In addition to selection interviews, human resource managers use other types of assessment methods for personnel selection, e.g., numerical or logical reasoning tests, personality tests, etc. Here, psychometric tests loosely refer to testing instruments that follow the psychological testing tradition, i.e., tests that have been subject to validation and standardization procedures, and are norm-referenced (see e.g., Anastasi, 1988 for a detailed discussion).

Common psychometric tests include general cognitive ability tests, job knowledge tests, and personality tests. According to Schmidt and Hunter's (1998) meta-analysis, the predictive validity of general cognitive ability tests (or general intelligence tests) for future job performance is .51. They further conclude that after 85 years of selection research, general cognitive ability assessment should be a primary selection method because of its high validity and low application cost. Other selection methods should be used as supplements to it.

In addition to intelligence tests, personality assessment is another type of commonly used psychometric test. The basic rationale of using personality assessment in personnel selection is that certain jobs would require individuals with certain personality characteristics, e.g., a good salesperson should be one who is outgoing and sociable. Thus, one major goal of personnel selection is to find the best match of personality orientation and job characteristics.

In reality, however, not all personality assessments are valid in predicting future job performance. According to some recent research on the relation between personality measures and job performance, some personality dimensions may be more promising than others. For instance, the predictive validity of integrity for future job performance was found to be .41 and that of conscientiousness was .31 (see Schmidt & Hunter, 1998). While these validity coefficients are not as high as those of cognitive ability tests or structured interviews, these personality measures can be considered as useful supplements for personnel selection.

It is also important to note that most of the personality inventories were developed in the West. Direct import of such instruments to the local population could be erroneous. Until recently, Cheung and her colleagues have developed the *Chinese Personality Assessment Inventory (CPAI)*, a comprehensive inventory measuring Chinese personality (Cheung et al, 1996; 2001). With such an instrument, it is now possible to measure the personality characteristics of our local employees in a reliable and valid manner.

C. Assessment Center

Since its introduction in the 1950s', the assessment center approach has become one common method for assessing managerial skills in organizations in the U.S. and Europe. The term of "assessment center" actually has nothing to do with a "center" or any fixed geographical location. It is used to describe a process by which an individual, or a group of individuals, is assessed by a team of judges (usually called "assessors") using a comprehensive and integrated series of assessment devices, many of which are performance tests.

Simply put, the development of an assessment center usually starts with a job analysis in order to identify the dimensions that are the important aspects of the job of interest. These job dimensions are then measured by various assessment center devices. Common job dimensions assessed include oral communication, planning and organization, work delegation, control, decisiveness, initiative, stress tolerance, adaptability, and tenacity (Gatewood & Feild, 1994). It is important to note that these dimensions can be defined in terms of actual

activities on the job.

In order to assess these dimensions, multiple devices are used. During a one- to two-day assessment, candidates will typically be asked to participate in various activities designed to capture the predetermined attributes or job dimensions. Common activities include in-basket task, group discussion, case analysis, simulation, personality assessment, and interview. For instance, in a simulation exercise, a job candidate may be asked to perform tasks that are typically performed by the job incumbent in a simulated setting. Candidates' performances will then be observed and scored by multiple assessors according to some preset criteria. The use of multiple assessors is mainly to ensure the reliability across raters (known as "inter-rater reliability"). Table 1 presents an example of how four different attributes are assessed by four different devices. For example, a candidate's oral communication skills will be assessed when he/she participates in the leaderless group discussion and the simulation exercise. It is generally believed that, with multiple assessment devices and assessors, the assessment center approach can serve as a more comprehensive and valid assessment method.

However, the assessment center approach also has its limitations. One commonly cited problem is that it is costly. For each selection, a certain degree of customization is required as different job positions would require different types of assessment devices. In addition, organizations generally have to rely on independent consulting firms to design and execute the entire assessment program because it requires a certain degree of technical knowledge (e.g., how to conduct the behavioral scoring objectively) and support (e.g., using confederates in simulation exercises). Furthermore, systematic research on assessment center is relatively lacking when compared to other commonly used assessment methods.

Table 1: An example of how different assessment devices examine various job skills

	Attributes Assessed			
Assessment Devices	Communi-cation Skill	Work Delegation	Taking Initiative	Planning & Organiza-Tion
In-basket Exercise		X		X
Leaderless Group Discussion	X		X	
Simulation	X	X	X	X
Case Analysis		X		X

III. General Recommendations

According to the above discussion, it is obvious that no one single assessment tool can address all the needs of employers. Thus, it is important to reiterate some important considerations when conducting a personnel selection exercise.

Prior to the selection exercise, a systematic job analysis should first be conducted to identity the core competencies that the candidate is expected to possess, regardless of the assessment method(s). These core competencies identified should then form the basis for the selection criteria. It should also be emphasized that core competencies usually consist of a broad spectrum of attributes. Therefore, it is sensible to use multiple assessment methods in order to increase the validity of the selection measures. Finally, as the above discussion is meant to be brief and introductory, readers are highly recommended to consult the following the references for further details.

IV References

Anastasi, A. (1988). *Psychological testing (8^{th} edition).* New York: Macmillan.

Brogden, H.E. (1949). When testing pays off. *Personnel Psychology, 2, 171-183*.

Cheung, F.M., Leung, K., Fan, R.M., Song, W.Z., Zhang, J.X., & Zhang, J.P. (1996). Development of the Chinese Persoanlity Assessment Inventory. *Journal of Cross-Cultural Psychology, 27(2), 181-199.*

Cheung, F.M., Leung, K., Zhang, J.X., Sun, H.F., Gan, Y.Q., Song, W.Z., & Xie, D. (2001). Indigenous Chinese personality construct: Is the Five-Factor Model complete? *Journal of Cross-Cultural Psychology, 32(4), 407-433.*

Gatewood, R.D., & Feild, H.S. (1994). *Human resource selection (3^{rd} edition).* New York: Harcourt Brace & Company.

Schmidt, F.L., Hunter, J.E., McKenzie, R.C., Muldow, T.W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology, 64, 609-626.*

Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124(2), 262-274*.

Sampling Design of the General Household Survey

Jonathan Ng
Census and Statistics Department

Introduction

The General Household Survey (GHS) is a sample survey which has been conducted by the Census and Statistics Department (C&SD) on a continuous month-to-month basis since August 1981. The main objective of the GHS is to collect data on the labour force, employment, unemployment and underemployment. Information on the demographic and socio-economic characteristics of the population is also collected. In addition, through the use of a supplementary part, the GHS also includes some specially designed questions to collect data on special topics required by other Government departments/policy bureaux. The survey makes use of scientific sampling design to ensure that the key labour force statistics compiled from the survey data are of high level of statistical precision. This paper briefly explains the various aspects of the sample design adopted in the GHS including the sampling frame, sampling unit, estimation method, etc.

Survey Coverage

The GHS covers the land-based non-institutional population (LBNP) of Hong Kong, with inmates of institutions and persons living on board vessels being excluded. The survey covers about 99% of the Hong Kong Resident Population.

Sampling Frame

The sampling frame of the GHS consists of two parts:

(i) Register of Quarters (RQ)

This contains computerized records of all addresses of permanent quarters in built-up areas, including urban areas, new towns and major developed areas. Each unit of quarters is identified by a unique address with details such as district (e.g. Tertiary Planning Unit and District Council District/Constituency Area), street name, building name, floor number and flat number. Besides, the information of the quarters such as the housing type and usage is included. Temporary structures (e.g. temporary structures on rooftop and in backlane) are not separately identified and are treated as attachments to the permanent quarters.

(ii) Register of Segments (RS)

This contains records of segments in non-built-up areas which are delineated by relatively permanent and identifiable landmarks such as footpath and river. There are about 10 quarters in each segment. The use of segments in constructing the sampling frame is necessary since the quarters in these areas may not have clear addresses and cannot be readily identified. All quarters in the sampled segments are listed and enumerated in field operation.

The frame of quarters is continuously updated using various administrative records on new buildings and demolitions and also by field verification of the C&SD staff. The boundary of segments would be redrawn every five years before the Population Census/By-census. The last update was conducted in end 1999 and early 2000.

Sampling Unit

The quarters in built-up areas and segments in non-built-up areas are used as the sampling unit in the survey. Only the quarters that are used for residential and partially residential purposes are covered by the survey. All households in the sampled quarters/segments are to be enumerated.

Sample Design

A rotational replicate sample design is adopted in the GHS. Relevant records in the frame of quarters are first stratified by geographical area and type of quarters (records in the RS are sorted by geographical area only). Replicates are then formed by drawing the quarters/segment records systematically according to a fixed sampling interval of around 1 in 4400. The entire frame of quarters is thus divided into some 4400 replicates, each containing about 500 quarters (including both permanent quarters and quarters in segments). Distribution of quarters/segment records by geographical area and type of housing in each replicate would generally follow the overall distribution pattern in Hong Kong. Each replicate is an independent representative sample of the target population.

In the GHS, 18 replicates are selected for enumeration in each month. Under the rotational replicate sample design, sampled quarters in half of these 18 replicates (i.e. 9 replicates) in the current month were enumerated (first enumeration) three months ago (old replicates). Another 9 replicates are newly drawn which are enumerated for the first time (new replicates) in the current month. The old replicates are then rotated out after the second enumeration. The sampled quarters in each replicate are thus enumerated twice at a three-month interval (i.e. at month t and month t+3). The rotational design can be depicted schematically as follows:

Quarter	Month	Replicates enumerated	No. of old replicates used	No. of new replicates used
First	Jan	X1 - X9*, A1 - A9	9	9
	Feb	Y1 - Y9*, B1 - B9	9	9
	Mar	Z1 - Z9*, C1 - C9	9	9
Second	Apr	A1 - A9, D1 - D9	9	9
	May	B1 - B9, E1 - E9	9	9
	Jun	C1 - C9, F1 - F9	9	9
Third	Jul	D1 - D9, G1 - G9	9	9
	Aug	E1 - E9, H1 - H9	9	9
	Sep	F1 - F9, I1 - I9	9	9

Note: * Replicates X1 - X9 were enumerated 3 months ago (i.e. in Oct of preceding year), while Y1 – Y9 and Z1 - Z9 were enumerated in Nov and Dec of preceding year respectively.

The overlapping of samples, coupled with rotational feature, leads to more precise estimation of changes over time on the one hand, and facilitates the control of respondents' burden on the other.

Sample Size

The monthly sample size of the GHS averages about 8 000 households. There are hence about 24 000 households in each three-month sample. The quarterly sampling fraction is about 1.2% of the target population. Currently, GHS achieves a response rate of about 90% involving about 22 000 successfully enumerated households in a quarter.

In order to ensure that the key statistics (e.g. unemployment rate) compiled from the survey data are of sufficient precision and at the same time to eliminate short-term fluctuations, data from samples for three consecutive months are pooled together so as to yield sufficiently precise statistics for release to the public. In other words, statistics are published on a moving three-month basis (e.g. Jan – Mar, Feb – Apr, Mar – May, …) rather than on a monthly basis.

Estimation Method

In deriving the estimates pertaining to the whole population, grossing-up factors are applied to the sampled observations with reference to the selection probability and response rates. Grossing-up factors for sampled persons in specific sex-age groups are obtained by ratio estimation using the sex and age distribution of the independent estimate of LBNP. In deriving the grossing-up factors, adjustment has been made for the differential response rates between the households in public and private sector housing. Estimates for the whole population are obtained by summing up individual estimates in respect of each specific sex-age group. An illustration of deriving the estimate of unemployment rate (UR) compiled from the GHS is given below:

Estimate the number of persons in specific sex-age group in specific housing type (i.e. public or private housing)

$$P_{ij} = \left(q_{ij} \times \frac{1}{r_j} \right) \times \frac{p_i}{\left(q_{ij'} \times \frac{1}{r_{j'}} \right) + \left(q_{ij''} \times \frac{1}{r_{j''}} \right)}$$

where

P_{ij} – estimated number of persons in the i^{th} sex-age group in housing type j

j = j' (public housing) or j"(private housing)

P_i – independent estimate of LBNP in the i^{th} sex-age group

$q_{ij'}$ and $q_{ij''}$ – numbers of persons in public and private housing respectively in the i^{th} sex-age group as obtained from GHS (sample count)

$r_{j'}$ *and* $r_{j''}$ – response rates of households in public and private housing respectively as obtained from GHS (sample count)

Grossing-up factor for a sampled person in housing type j in the ith sex-age group:

$$GU_{ij} = \frac{P_{ij}}{p_{ij}}$$

where

GU_{ij} – grossing up factor for sampled persons in housing type j in the i^{th} sex-age group

P_{ij} – estimated number of persons in housing type j in the i^{th} sex-age group

p_{ij} – sample count of persons in housing type j in the i^{th} sex-age group

Using the grossing-up factors for sampled persons, the UR is estimated as

$$UR = \frac{\sum u_{ij} \times GU_{ij}}{\sum l_{ij} \times GU_{ij}}$$

where

GU_{ij} – grossing up factor for sampled persons in housing type j in the i^{th} sex-age group

l_{ij} – sample count of labour force in housing type j in the i^{th} sex-age group

U_{ij} – sample count of unemployed persons in housing type j in the i^{th} sex-age group

Variance Estimation Method

In view of the complicated procedures in estimation and the replicate sample design of the GHS, Jacknife Repeated Replication (JRR) method is used for estimating the variance for the GHS statistics. Briefly, JRR method estimates the variance based on the variability in the estimated parameter of interest between the subsamples which are formed by systematically omitting some observations in the parent sample. For the variance of estimated UR, it is obtained by systematically omitting one replicate (there are 54 replicates in a quarterly sample) each time and using the data of the remaining 53 replicates for computation. The process repeats 54 times until all replicates have been omitted once. As a result, there will be 54 pseudo estimates. The variance is calculated as below:

$$\text{var}(\hat{UR}) = \frac{(1-f)(k-1)}{k}\sum_{k=1}^{54}(UR_k - \hat{UR})^2$$

where

UR_k - pseudo estimate for unemployment rate compiled after systematically omitting a replicate each time (i.e. based on data in the remaining 53 replicates). That is the estimate computed from reduced sample by omitting the k^{th} replicate. The process repeats k times until all replicates have been omitted once.

$\hat{UR}$ - the estimate for unemployment rate compiled from al1 replicates in a quarterly sample

On the whole, JRR method provides a simple and convenient method for estimating the variance of the GHS statistics.

Book Review

The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century By David Salsburg

Ping-shing Chan
The Chinese University of Hong Kong

This book has a very whimsical title. What is the relationship between Lady Tasting Tea and Statistics? In fact, it refers to a Cambridge University tea party at which a lady insisted, *Tea tasted different depending upon whether the tea was poured into the milk or whether the milk was poured into the tea*. It seems nonsense but R. A. Fisher immediately devised a careful experiment that was largely capable of ruling out the effect of random luck. The result of the experiment is that the lady correctly identified each cup.

This is the first story given in the book. The author then describes the life and work of many statistics masters including Francis Galton, Karl Pearson, Egon Pearson, Jerzy Neyman, Abraham Wald, John Tukey, E.J.G. Pitman, Ed Deming, R.A. Fisher, George Box, David Cox, Gertrude Cox, Emil Gumbel, L.H.C. Tippett, Stella Cunliffe, Florence Nightingale Daivd, William Sealy (Student) Gosset, Frank Wilcoxon, I.J. Good, Harold Hotelling, Morris Hansen, William Cochran, Persi Diaconis, Brad Efron, PaulLevy, Jerry Cornfield, Samuel Wilks, Andrei Kolmogorov, Guido Castelnuovo, Francesco Cantelli and Chester Bliss.

Through the descriptions, Dr. Salsburg examines the development of Statistics in the early 20th century, which is when the Statistical Revolution took place. Important developments are clearly explained in their historical context, and their implications for current scientific research are given.

The targeted audience of this book is the general public. You will find no mathematical symbols and formulas in this book. However, you will find a lot of important discussion of current issues of statistics such as role of epidemiology, use of martingale methods in survival analysis, exploratory data analysis, p-values, Bayesian models, non-parametric methods, bootstrap, hypothesis tests and confidence intervals. It is a great deal of value in reviewing the key ideas and philosophy of the great statisticians of the 20th century.

The book includes a timeline that provides a list in chronological order of important events and the associated

personalities in the history of statistics.

This is a readable, enjoyable book that shows the impact of statistics and statistical thought on science.

News

Department of Statistics and Actuarial Science, The Hong Kong University

Prof. H. Tong awarded Distinguished Research Achievement Award of Hong Kong University.

Prof. W.K. Li awarded Outstanding Researcher Award of Hong Kong University.

Prof. Stephen Lee awarded Outstanding Young Researcher Award of Hong Kong University.

Department of Statistics, The Chinese University of Hong Kong

Prof. Chan Ngai Hang awarded a Fellow of Institute of Mathematical Statistics.