UTILITY ANALYSIS IN STATISTICS AND APPLIED PSYCHOLOGY

Jifeng Wang
Department of Mathematical Sciences
Montana State University

January 9, 2003

A writing project submitted in partial fulfillment of the requirements for the degree

Master of Sciences in Statistics

APPROVAL

of a writing project submitted by

JIFENG WANG

This writing project has been read by the writing project director and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Statistics Faculty.

January 27, 2003
Date

James Robison Cox
James F. Robison-Cox
Writing Project Director

Contents

INTRODUCTION AND PREVIEW

In an uncertain and complex world, the decision maker must balance judgments about uncertainties and complexities with his or her preferences for possible consequences or outcomes. It is not easy to do. Here we suggest utility analysis which concentrates on formalizing the preference or value side of the problem. The major contributions to the growth of utility analysis have come out of economics, statistics, mathematics, psychology, and the management sciences (Fishburn, 1970, p. vii). We would focus on two fields – statistics and applied psychology in this writing project.

The power of the concept of utility and the grounds for the interest in this writing project is as follows. If an appropriate utility is assigned to each possible consequence and the expected utility of each alternative is calculated, then the best course of action is the alternative with the highest expected utility.

In psychology, the term utility is used to describe and compare the economics of a personnel strategy to an organization. The rationale is that if a company has higher performing employees, it will make more money. Utility loss occurs if the organization departs from the optimum strategy of promoting the most qualified candidates. We substantiate this by citing a practical example – "glass ceiling", which refers to the underrepresentation of women in senior levels of organizations. To assess the impact of small amounts of gender bias in subjective performance ratings on women's organizational mobility, we will use computer simulation as a method.

PART I – UTILITY ANALYSIS IN STATISTICS

1.1 Utility

We all make decisions all the time. In evaluating the consequences of possible actions, we may encounter two major problems. The first is that the values of the consequences may not have any obvious scale of measurement. For example, it is not clear how to evaluate the importance of the reputation to businesses in a concrete way. Secondly, even when there is a clear scale by which consequences can be evaluated, the scale may not reflect "true" value to the decision maker. For instance, a millionaire will most probably not bother to do an unpleasant task that will earn him $100. But a penniless person might well value the 100 dollars enough to do the task. In other words, the value of $1,000,100 is probably not the same as the value of $1,000,000 plus the value of $100.

To work mathematically with ideas of "value", it will be necessary to assign numbers indicating how much something is valued. Such numbers are called utilities (Berger, 1985, p. 47). The general conceptual definition of utility is the quality or state of being useful.

Utility appears to have been first discussed by Daniel Bernoulli in *1738* (Watson, 1987, p. 19). He suggested that the worth, or utility, of money for each individual was non-linear and had a decreasing slope; marginal utility decreased as wealth increased. Bernoulli's concern was to explain the St Petersburg paradox, which involves the proper evaluation of a particular uncertain gamble (Watson, 1987, p. 19-20). When there is a small probability of winning a large prize, people are not willing to pay very much money to play the game.

1.2 Utility Theory

We will mainly refer to James Berger "Statistical Decision Theory and Bayesian Analysis" (Second Edition) to introduce utility theory, the construction of utility function, and the utility of money.

It is necessary to define some key words first. The set of all consequences of interest will be called the set of rewards, and will be denoted by $\mathcal{R}$. Often the elements of $\mathcal{R}$ will consist of nonnumerical quantities.

For example, you face the decision of what to do after graduation from university. There are two choices available: a_1 - apply for a job, and a_2 - take an offer to be a graduate student. You would really rather apply for a job but there is a 40% chance of being refused. Let θ_1 denote the state of nature "being refused by the employer" and θ_2 denote "being accepted by the employer". Clearly $\pi(\theta_1) = 0.4 = 1 - \pi(\theta_2)$. The four consequences of rewards of interest are $r_1 = (a_1, \theta_1), r_2 = (a_1, \theta_2), r_3 = (a_2, \theta_1)$, $r_4 = (a_2, \theta_2)$. Here (a_i, θ_j) denotes action being taken and the state of nature that occurs. The preference ordering is:

$$r_1 \prec r_4 \prec r_3 \prec r_2$$

The reasoning is that r_1 is the worst, since you apply for a job rather than taking the offer to be a graduate student, but you are refused; r_4 is okay, but you will regret somewhat not having applied for a job since you would be employed; r_3 is good, and you congratulate yourself on making a good decision about your future; and r_2 is the best, since you applied for a job and were accepted – you get what you really wanted.

Since we are uncertain as to which of the possible consequences will actually occur, the results of actions have frequency probability distributions on $\mathcal{R}$. Let $\mathcal{P}$ denote the set of all such probability distributions. To work with values and preferences concerning probability distributions in $\mathcal{P}$, we would prefer a real-valued function $U(r)$ to be constructed such that the "value" of a probability distribution $P \in \mathcal{P}$ would be given by the expected utility $E^P[U(r)]$. This real-valued function is called a utility function if it exists. The definition of Utility Theory from *Encyclopedia of Statistical Sciences (1985)* is: "Utility theory is the study of quantitative representations of people's preferences and choices".

To start the construction of U, we have one assumption that it is possible to state preferences among elements of $\mathcal{P}$. The following notations will be used:

If P_1 and P_2 are in $\mathcal{P}$,

$P_1 \prec P_2$ indicates that P_2 is preferred to P_1,

$P_1 \approx P_2$ indicates that P_1 is equivalent to P_2,

$P_1 \prec= P_2$ indicates that P_1 is not preferred to P_2.

The goal is to find a function $U(r)$ such that if P_1 and P_2 are in $\mathcal{P}$, then P_2 is preferred to P_1 if and only if $E^{P_1}[U(r)] < E^{P_2}[U(r)]$. The function U is then the desired quantification of the decision maker's preference or value pattern, and can be called a utility function.

1.3 Construction of Utility Function

Step 1. Choose two rewards, r_1 and r_2, which are not equivalent. Assume they are labeled so that $r_1 \prec r_2$. Let $U(r_1)=0$ and $U(r_2)=1$. It is best to choose r_1 and r_2 in a convenient fashion. The choices of r_1 and r_2 really serve only to set scale for U. However, values of $U(r)$ for other r will be established by comparison with r_1 and r_2. So r_1 and r_2 should be chosen to make comparisons as easy as possible.

Step 2. For a reward r_3 such that $r_1 \prec r_3 \prec r_2$, find the α $(0<\alpha<1)$ such that

$$r_3 \approx P = \alpha\langle r_1\rangle + (1-\alpha)\langle r_2\rangle$$

Notation $\langle r\rangle$: A reward $r \in \mathcal{R}$ will be identified with the probability distribution in $\mathcal{P}$ which gives probability one to the point r. This probability distribution will be denoted $\langle r\rangle$.

Define

$$U(r_3) = E^P[U(r)] = \alpha U(r_1) + (1-\alpha)U(r_2)$$

$$U(r_3) = 1-\alpha$$

because $U(r_1)=0, U(r_2)=1$

Since r_3 is preferred to r_1, r_2 is preferred to r_3, there should be a "gamble", in which you get r_1 with probability α and r_2 with probability $(1-\alpha)$, which is of the same value as r_3.

Step 3. For a reward r_3 such that $r_3 \prec r_1$, find the α $(0<\alpha<1)$ such that

$$r_1 \approx P = \alpha\langle r_3\rangle + (1-\alpha)\langle r_2\rangle$$

Then to have

$$U(r_1) = E^P[U(r)] = \alpha U(r_3) + (1-\alpha)U(r_2)$$

$$0 = \alpha U(r_3) + (1-\alpha)$$

$$U(r_3) = \frac{-(1-\alpha)}{\alpha}$$

Step 4. For a reward r_3 such that $r_2 \prec r_3$, find the α $(0<\alpha<1)$ such that

$$r_2 \approx P = \alpha\langle r_1\rangle + (1-\alpha)\langle r_3\rangle$$

Then to have

$$U(r_2) = E^P[U(r)] = \alpha U(r_1) + (1-\alpha)U(r_3)$$
$$1 = (1-\alpha)U(r_3)$$
$$U(r_3) = \frac{1}{1-\alpha}$$

Step 5. Periodically check the construction process for consistency by comparing new combinations of rewards. For example, assume the utilities of r_3, r_4 and r_5 have been found by the preceding technique, and that $r_3 \prec r_4 \prec r_5$. Then find the α $(0 < \alpha < 1)$ such that

$$r_4 \approx P = \alpha\langle r_3\rangle + (1-\alpha)\langle r_5\rangle$$

It should be true that

$$U(r_4) = E^P[U(r)] = \alpha U(r_3) + (1-\alpha)U(r_5)$$

If this relationship is not (approximately) satisfied by the previously determined utilities, then an error has been made and the utilities must be altered to attain consistency.

We will go back to the example of what to do after graduation. As mentioned in "Utility Function", the four consequences of rewards of interest are $r_1 = (a_1, \theta_1)$, $r_2 = (a_1, \theta_2)$, $r_3 = (a_2, \theta_1)$, $r_4 = (a_2, \theta_2)$. The preference ordering is

$$r_1 \prec r_4 \prec r_3 \prec r_2$$

To construct U, it is natural to use the worst and best rewards, r_1 and r_2, as the initial points. Hence let $U(r_1) = 0$ and $U(r_2) = 1$. To determine $U(r_4)$, compare r_4 with the gamble $\alpha\langle r_1\rangle + (1-\alpha)\langle r_2\rangle$. Assume it is concluded that r_4 is equivalent to $0.4\langle r_1\rangle + 0.6\langle r_2\rangle$. In other words, you would just as soon take the offer to be a graduate student and get the offer from the company to whom you applied a job, as you would play the gamble of r_1 with probability 0.4 or the best decision r_2 with probability 0.6. Thus $U(r_4) = (1-\alpha) = 0.6$. Likewise, let us say it is determined that

$$r_3 \approx 0.3\langle r_1\rangle + 0.7\langle r_2\rangle$$

Then $U(r_3) = 1 - \alpha = 0.7$.

At this point, we need to check the construction of U by comparing a different combination of rewards, say r_2, r_3, and r_4. Find an α such that

$$r_3 \approx \alpha\langle r_4\rangle + (1-\alpha)\langle r_2\rangle$$

Let us say α=0.6 is felt to be correct. But if the utility function is accurate, we know that

$$U(r_3) = E^P[U(r)] = \alpha U(r_4) + (1-\alpha)U(r_2)$$
$$0.7 = 0.6\alpha + (1-\alpha)$$

because as we stated above $U(r_3) = 0.7, U(r_4) = 0.6,$ and $U(r_2) = 1$

Solve for α: $\alpha = 0.75$.

This does not agree with the value of 0.6 that was obtained from direct comparison. Hence there is an inconsistency. It is thus necessary to go back and re-examine all the comparisons made, until a consistent set of utilities is obtained. Let us assume that the end result of this process is that $U(r_4) = 0.6$ and $U(r_3) = 0.75$ are decided upon.

Having the utility function, we can proceed directly to find the optimal action in terms of expected utility. The expected utility is the proper way to evaluate uncertain rewards. Clearly, the expected utility of action a_1 is

$$\begin{aligned} E^\pi[U(r)] &= \pi(\theta_1)U(r_1) + \pi(\theta_2)U(r_2) \\ &= (0.4)(0) + (0.6)(1) \\ &= 0.6 \end{aligned}$$

While the expected utility of action a_2 is

$$\begin{aligned} E^\pi[U(r)] &= \pi(\theta_1)U(r_3) + \pi(\theta_2)U(r_4) \\ &= (0.4)(0.75) + (0.6)(0.6) \\ &= 0.66 \end{aligned}$$

Thus the optimal action is to take the offer to be a graduate student.

1.4 The Utility of Money

As indicated by the example of $1,000,000 and $100 in the introduction, the difference in value between $(z+100) and $z is decreasing as z increases. An additional $100 is quite valuable when z=0, but of little importance when z=1,000,000. That is to say, the marginal value of money for most people is decreasing.

To get a better feeling for such a utility function $U(r)$, let us consider the construction of a personal utility function for money. We may begin, as before, by choosing $r_1 < r_2$ and setting $U(r_1) = 0$, $U(r_2) = 1$. It is assumed that if $r_1 < r_2$, then $r_1 \prec r_2$.

Step 1. Find r_3 so that $U(r_3) = \frac{1}{2}$. Since, for $P = \frac{1}{2}\langle r_1 \rangle + \frac{1}{2}\langle r_2 \rangle$,

$$\frac{1}{2}U(r_1) + \frac{1}{2}U(r_2) = \frac{1}{2} = E^P[U(r)],$$

this is equivalent to finding the reward r_3 which is valued as highly as the gamble $\frac{1}{2}\langle r_1 \rangle + \frac{1}{2}\langle r_2 \rangle$.

Step 2. Find r_4 and r_5 such that $U(r_4) = \frac{1}{4}$ and $U(r_5) = \frac{3}{4}$. Since

$$\frac{1}{2}U(r_1) + \frac{1}{2}U(r_3) = \frac{1}{2}(0) + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4}, \text{ and}$$

$$\frac{1}{2}U(r_3) + \frac{1}{2}U(r_2) = \frac{1}{2}\left(\frac{1}{2}\right) + \frac{1}{2}(1) = \frac{3}{4}$$

r_4 is obviously the point equivalent to the gamble $\frac{1}{2}\langle r_1 \rangle + \frac{1}{2}\langle r_3 \rangle$ and r_5 is the point equivalent to $\frac{1}{2}\langle r_3 \rangle + \frac{1}{2}\langle r_2 \rangle$.

Step 3. Since

$$U(r_3) = \frac{1}{2}U(r_4) + \frac{1}{2}U(r_5),$$

r_3 must be equivalent to the gamble $\frac{1}{2}\langle r_4 \rangle + \frac{1}{2}\langle r_5 \rangle$. This provides a valuable check on the determinations made so far.

Step 4. Continue the above process of finding the points with utilities $i/2^n$ and checking for consistency, until a sufficient number of points have been found to enable a graph of $U(r)$ to be made. Indeed the function typically levels off, i.e., is bounded.

The advantages of this method of finding $U(r)$ are that only simple gambles of the form $\frac{1}{2}\langle r_i \rangle + \frac{1}{2}\langle r_j \rangle$ need to be considered, and also that it is easy to check for consistency.

PART II -- UTILITY ANALYSIS IN APPLIED PSYCHOLOGY

2.1 Utility in Psychology

In statistics, utility is a numerical measure for measuring the "gain" accruing from any combination of true parameter and decision. Utility is measured on such a scale that the objective can be taken to be to maximize expected utility (Cox & Hinkley, 1974, p. 412). In psychology, the term utility is used to describe and compare the economics of a personnel strategy to an organization, including the economic positive advantages and the negative economic consequences of a strategy. The rationale is that if a company has higher performing employees, it will make more money.

Clark Hull (1928) suggested that utility might be gained by moving the mean of the performance distribution up toward the high end of the performance scale, that is, by hiring the top $p\%$ of performers. To meet the need for a common metric for indexing job performance, Brogden and Taylor (1950) suggested a dollar metric – that the output of many workers be measured and then converted to a dollar metric, from which the standard deviation of the resulting distribution could be computed. They proposed to use traditional cost-accounting methods to calculate the worth of performance in dollar terms. Schmidt, Hunter, McKenzie, and Muldrow (1976) developed a method of estimating the dollar value of jobs that was more direct than the Brogden cost-accounting procedures and that met the demands of a ratio scale of measurement. Cronbach and Gleser (1965) introduced similar interaction models in their discussion of the utility of differential placement. They proposed that the overall utility of a personnel decision system could be improved by sequencing selection decisions and using appropriate cost-benefit decision rules. The Brodgen-Cronbach-Gleser (BCG; Cronbach, 1965) utility model is a framework for determining the economic impact of changes in certain human resource practices on organizational performance (Schmidt & Rauschenberger, 1986). While the BCG model can be used to estimate the impact of a wide range of human resource practices, it is most often considered in the context of employment selection.

2.2 Basic equation of utility model in applied psychology

Assume employees have a performance score, X, and that higher performance scores mean that the company makes more money. The selection procedures an organization uses determine in large part the performance capabilities of its workforce and therefore, determine organizational productivity. In making the decision whether to use a given selection procedure, it is critical that management have reliable estimates of the economic impact of using or not using the procedure. What we will present here is the information that is needed to calculate the expected productivity impact of selection procedures that are available to organizations.

The basic equation is:

$$\Delta U = \Delta r_{zy} * SD_y * \overline{Z}_x * T * N_h - \Delta C * N_a$$

(Schmidt, F. L. & Rauschenberger, J. 1986)

where:

y is defined to be the productivity in dollars.

ΔU is the gain in productivity in dollars as a result of using the new selection procedure.

Δr_{zy} is the increase in the correlation between selection procedure scores and later job performance. The increase results from improved selection methods. Note that if the old selection method has no validity (i.e., r_{zyold}=0) then Δr_{zy} is equal to the validity of the new selection method; in this case, Δr_{zy} is equal to the correlation between scores on the new selection method and later job performance. Otherwise, Δr_{zy} is equal to the difference between the correlations (validities) for the old and new selection methods.

SD_y is the standard deviation of the dollar value of yearly output of randomly selected employees. In the past this has been the most difficult and expensive item of required information to estimate. However, Hunter and Schmidt (1982) have shown that SD_y can be conservatively estimated at 40% of annual salary. We could determine SD_y by computing the average annual salary of new selectees in the job and take 40% of that value.

$\overline{Z}_{x_h}$ is the average selection procedure score for those hired in standard score form. Scores are standardized on the applicant pool. We could compute a z score for each person hired and take the average of these to get the value of $\overline{Z}_{x_h}$. The z-scores are calculated using the formula (applicant score – mean score of all applicants) / *SD* of all scores. The average z score for all applicants will be zero by definition. The relationship between $\overline{Z}_{x_h}$ and *p*, which is the selection ratio, will be discussed soon.

T is the average number of years that all individuals selected for this job remain in that position.

N_h is the number of individuals who are hired using the new selection procedure.

ΔC is the change in cost of administering the selection mechanism to one person. If no selection method had been used previously, *C* is the full cost of implementing the procedure per person tested.

N_a is the number of applicants who apply for the position in question.

2.3 Deriving the Basic equation of utility model

We will use the principles of linear regression to demonstrate how the selection ratio (*p*) and the standard deviation of job performance in dollars (SD_y) affect the economic utility of a selection device.

The basic linear model is:

$$Y = \beta Z_x + \mu_a + \varepsilon$$

where Y = job performance measured in dollar value; β = the linear regression weight on test scores for predicting job performance; Z_x = test performance in standard score form in the applicant group; μ_a = mean job performance (in dollars) of all applicants; and ε = individual error.

The estimate of β is:

$$\hat{\beta} = r_{zy} \frac{SD_y}{SD_z} = r_{zy} SD_y$$

where r_{zy} = the correlation between the test and job performance measured in dollar value; SD_y = the standard deviation of job performance measured in dollar value among applicants; SD_Z = the standard deviation of standard job performance among randomly selected employees. $SD_Z = 1$.

The equation that gives the average job performance is:

$$E(Y) = E(\beta Z_{x_h}) + E(\mu_a) + E(\varepsilon)$$
$$E(Y) = \beta E(Z_{x_h}) + \mu_a$$

Using method of moments estimators, this becomes:

$$\overline{Y}_h = \hat{\beta}\overline{Z}_{x_h} + \overline{y}_a$$
$$\overline{Y}_h = r_{zy} S_y \overline{Z}_{x_h} + \overline{y}_a$$
$$\overline{Y}_h - \overline{y}_a = r_{zy} SD_y \overline{Z}_{x_h}$$

The above equation gives mean gain in productivity per selectee resulting from use of the test, that is,

$$\Delta\overline{U} / selectee = r_{zy} * SD_y * \overline{Z}_{x_h}$$

where $\Delta\overline{U}$ is the marginal utility.

If we consider N_h, the number of people hired, the total utility gain is:

$$\Delta U / year = r_{zy} * SD_y * \overline{Z}_{x_h} * N_h$$

If we consider *T*, the average numbers of years that all individuals hired for this job remain in that position, then

$$\Delta U = r_{zy} * SD_y * \overline{Z}_{x_h} * T * N_h$$

We also need to consider the cost of testing one applicant C, which should be substracted from the total utility gain:

$$\Delta U = r_{xy} * SD_y * Z_x * T * N_h - C * N_a$$

where N_a = number of applicants.

The case we considered here is deciding whether to begin using a selection procedure where none is now used. To make the basic equation applicable to two more cases, i.e., deciding whether to replace one selection procedure with another that might be preferable, and deciding whether to add new measures to existing selection measures, we can make the basic equation we derived above more general by replacing r_{zy} by Δr_{zy}, and C by ΔC. Thus the basic equation becomes:

$$\Delta U = \Delta r_{zy} * SD_y * Z_x * T * N_h - \Delta C * N_a$$

2.4 Relationship between the selection ratio and average test score of those hired

We assume that the relation between the test and job performance is linear. We further assume that the test scores are normally distributed. In a normal distribution, the average test score of those hired $\overline{Z}_{x_h}$ can be computed from the selection ratio *p*, which is the percentage of applicants who are hired $(= \frac{N_h}{N_a})$. $\overline{Z}_{x_h} = \frac{\phi}{p}$, where ϕ is the ordinate in *N*(0,1) at the point of cut corresponding to *p*. The relationship between *p* and $\overline{Z}_{x_h}$ is discussed as follows.

$$\text{when } p \geq 0.5, \phi = \int_{\Phi^{-1}(p)}^{\infty} u\Phi(u)du$$

$$\text{when } p < 0.5, \phi = \int_{\Phi^{-1}(1-p)}^{\infty} u\Phi(u)du$$

Let $q = \Phi^{-1}(\max(p, 1-p))$, ϕ can be expressed as one equation:

$$\phi = \int_q^{\infty} u\Phi(u)du$$

$$= \int_q^{\infty} u \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Let $w = u^2/2$

$dw = udu$

when $u = q, w = q^2/2$

$$\phi = \frac{1}{\sqrt{2\pi}} \int_{q^2/2}^{\infty} e^{-w} dw$$

$$= -\frac{1}{\sqrt{2\pi}} e^{-w} \Big|_{q^2/2}^{\infty}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-q^2/2}$$

$$\overline{Z}_{x_h} = \frac{\phi}{p} = \frac{\frac{1}{\sqrt{2\pi}} e^{-q^2/2}}{p}$$

We can use R to simplify the calculations. For example, when p=0.8 and 0.3, the R-code to calculate $\overline{Z}_{x_h}$ and the results are:

```
> p<-0.8
> q<-qnorm(max(p,1-p))
> Z<-((1/sqrt(2*pi))*exp(-q^2/2))/p
> Z
[1] 0.3499524
> round(Z,2)
[1] 0.35

> p<-0.2
> q<-qnorm(max(p,1-p))
> Z<-((1/sqrt(2*pi))*exp(-q^2/2))/p
> Z
[1] 1.399810
> round(Z,2)
[1] 1.4
```

In this way we may get $\overline{Z}_{x_h}$ values corresponding to different p values automatically in R.

2.5 Variations to the basic equation of utility model

As we mentioned before, there are three basic variations to this equation. They are:

1. Deciding whether to begin using a selection procedure where none is now used.
2. Deciding whether to replace one selection procedure with another that might be preferable.
3. Deciding whether to add new measures to existing selection measures.

Here we would discuss case 2.

2.5.1 Case 2 study – Deciding whether to replace one selection procedure with another that might be preferable

In some cases, an employer must decide whether to replace one selection procedure with another that is either more valid or that costs less. This occurs when the employer, for one reason or another, cannot (or will not) use both procedures together.

Since we are comparing the validity of a new process with an old process, both the validity and cost are comparisons against previous values that are not zero. Thus Δr_{zy} becomes $r_{zynew} - r_{zyold}$, and ΔC becomes $C_{new} - C_{old}$. The equation for case 2 is:

$$\Delta U = (r_{zynew} - r_{zyold}) * SD_y * Z_{x_h} * T * N_h - (C_{new} - C_{old}) * N_a$$

For example, if we assume the old selection procedure being replaced had a validity of 0.30 and a cost per applicant of $50. Suppose the value of the validity obtained from the consultant's report for the new selection procedure is 0.50. Suppose also that the average yearly salary for new employees in this job is $25,000. Then SD_y is estimated as 25,000*0.40=$10,000. The management plans to select the top scoring 30 applicants who represent 30% of all applicants. Then the selection ratio is 0.30. The corresponding *Zx* value from the table is 1.17. Suppose it costs $65 per applicant to use the new selection procedure. Finally, you estimate that the new hires in this position will stay with your organization for an average of 4 years.

The estimates of the parameters based on the information above are:

r_{zynew}=0.50, r_{zyold}=0.30, SD_y=$10,000, p=0.30, T=4, N_h=30, C_{new}=$65, C_{old}=50, N_a=100

Using the R code, we get the Z_{x_h} value when p=0.30:

```
> p<-0.3
> q<-qnorm(max(p,1-p))
> Z<-((1/sqrt(2*pi))*exp(-q^2/2))/p
> round(Z,2)
[1] 1.16
```

Thus Z_{x_h}=1.16.

Then ΔU would be calculated as follows:

$$\Delta U = (r_{xynew} - r_{xyold}) * SD_y * Z_{x_h} * T * N_h - (C_{new} - C_{old}) * N_a$$
$$= (0.50 - 0.30) * 10{,}000 * 1.17 * 4 * 30 - (65 - 50) * 100$$
$$= 278{,}400 - 1500$$
$$= 276{,}900$$

In this example ΔU is positive, which indicates that the company will have utility gain if switch to the new selection procedure.

We would talk about another example of case 2, "glass ceiling", in details as follows.

2.5.2 "Glass Ceiling"

"Glass ceiling" refers to the underrepresentation of women in senior levels of organizations. Currently, less than 5% of senior-level management positions in Fortune 500 companies are held by women, a percentage that has changed little over the past twenty years (Catalyst, 1998). We would think of gender bias that can seriously disadvantage managerial women. Decision on promotion may give men an unearned performance increase, or bias, so that men with lower performance get promoted above women who actually score higher than the men. There is a corresponding economic cost to the organization, often referred to as a "utility loss", if the organization departs from the optimum strategy of promoting the most qualified candidates. This loss can be measured in dollars. Utility analysis is used in this example to estimate the cost of gender bias in performance evaluations to a business organization.

The equation for change in utility under different promotional strategies for one year at one level of rank is slightly different from the equation of case 2. The costs are assumed equal for one year. The equation is expressed as:

$$\Delta U = r_{xy} * \left(\frac{\lambda}{SR}\right) * SD_y * (\bar{x}_1 - \bar{x}_2) * \mathrm{N}$$

(Martell, R. F. & Robison-Cox, J. F., 2002)

where r_{xy} is the validity of performance score, x, associated with dollar value; SDy is the standard deviation of dollar value; $\bar{x}_1$ and $\bar{x}_2$ are average performance scores under two different promotional strategies; N is the number of employees involved; SR is the selection ratio and λ the performance cutoff score. Note that it is the upper levels of management that we are most concerned with as women managers are most excluded at this level.

Our goal is to find the average loss in utility due to giving men an unearned bias. The different promotional strategies we would use include the promotional strategy without bias to women, and the promotional strategies with bias to women at different levels. If a strong link between gender bias and dollar value can be shown, senior-level decision makers in organizations will have a business-related reason to aggressively root out sources of gender bias in personnel actions.

Because of the complex interactions among individuals in the organization and the mix of organizational factors that regulate an individual's movement up the organizational hierarchy, we would use computer simulation as a method for assessing the impact of gender bias in organizations. We would refer to Dr. Jim Robison-Cox's simulation method in R to conduct it.

PART III -- SIMULATION

3.1 Mean performance and weighted average performance scores for women and men at different gender bias level

Some new parameters will be added, among which the most important is r^2, the proportion of variance in performance scores to be explained by gender bias. Three r^2 values will be applied:

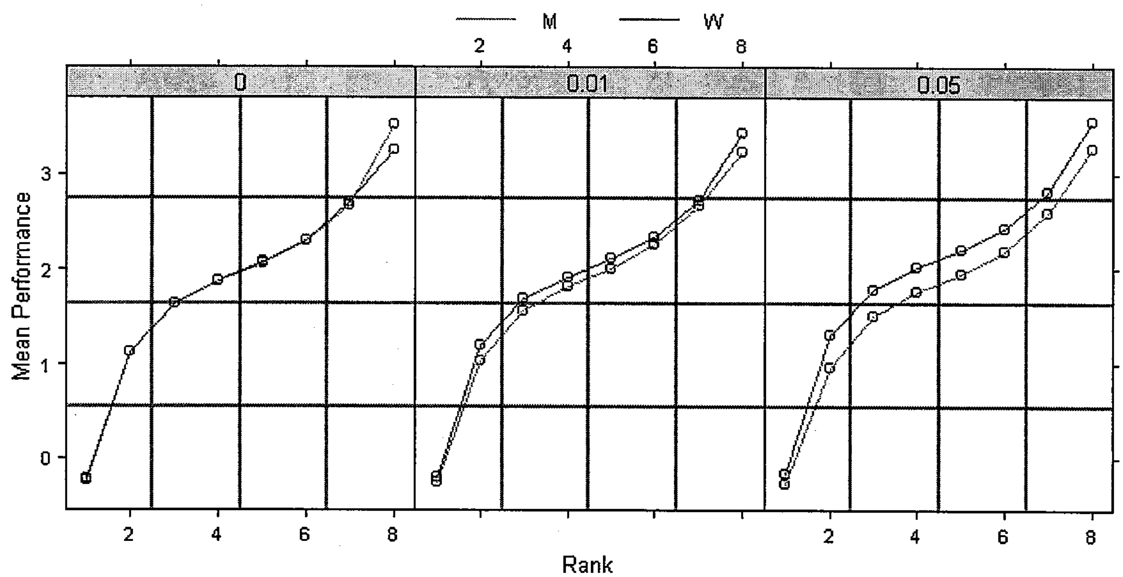
r^2=0 (no bias),

r^2=1% (gender bias effects account for 1% of the variability in subjective performance ratings), and

r^2=5% (gender bias effects account for 5% of the variability in subjective performance ratings).

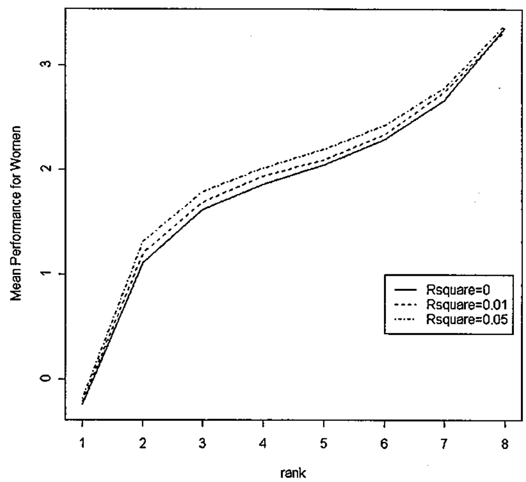
We would envision the organization as a data matrix wherein each row corresponds to an individual and columns are the characteristics, such as gender, performance score, rank, time-in-rank, and age. In the simulation we have access to true performance scores for all employees (gender bias is removed to get the true scores), so the $\bar{x}$'s in the utility equation are known exactly. Suppose we have 8 ranks with 1 the lowest and 8 the highest for the upper levels of management in one company.

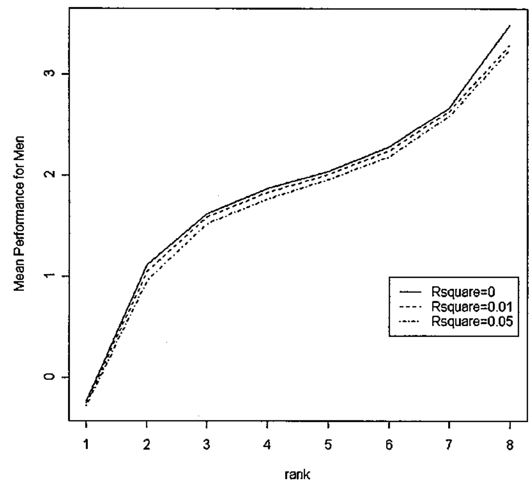
We would first look at the mean performances for men and women at each of the 8 ranks when r^2=0, r^2=1% and r^2=5% respectively. The graphs for the mean performances at three different situations are as follows:



It is obvious that as r^2 gets increased, the mean performances for women get higher and higher than the mean performances for men at each rank especially at ranks from rank 2 to 7. This means that women tend to have higher mean performance than men who are at the same rank as r^2 increases. This will make us think that decisions on promotion may give men an unearned performance increase, or bias, so that men with lower performance get promoted above women who actually score higher than the men.

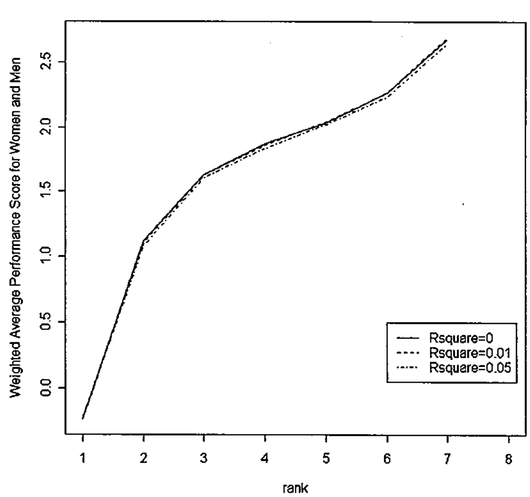
We may view the same issue at another angle by using the separate plots of mean performance for women and men at three different levels of r^2.





The mean performance for women gets increased at each rank except the highest rank when the bias increased. The mean performance for men gets decreased at each rank except the highest rank when the bias increased. There is an exception for rank 8 because we assume that there is only one person for this highest position, which is usually CEO, in the company.

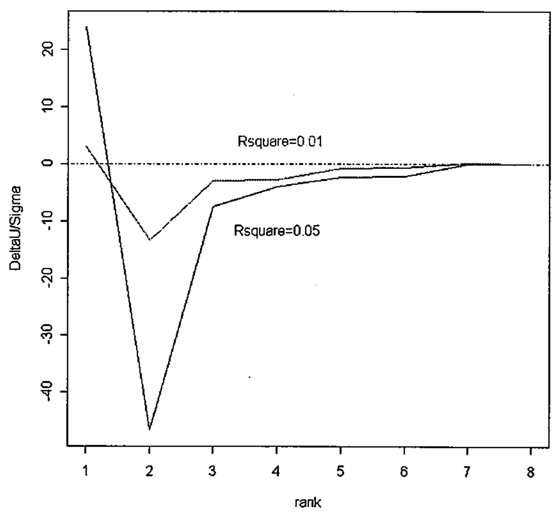
By simulation we may get the gender percentage for each rank at different r^2 levels. Using the gender percentage, we may get the weighted average performance score for women and men at each rank, which is: gender percentage of women*mean performance of women + gender percentage of men*mean performance of men at a specific rank. The plot for the weighted average performance score for women and men at all ranks at different r^2 levels is as follows:



It seems that the weighted average performance score for women and men are pretty close at different r^2 levels. But there is little difference among the three curves. When $r^2 = 0.05$, the weighted average performance score is about the lowest. The difference is not big, though.

3.2 Utility loss due to gender bias

What is most of interest to us, or to the business, is the average loss in utility, which is the reverse of utility gain, due to giving men an unearned bias. We assume that utility to the company is proportional to performance. The average loss in utility is denoted as $\Delta U / \sigma$. The combined graph for $\Delta U / \sigma$ when r^2=0, r^2=1% and r^2=5% are as follows:



The line y=0 indicates there is no utility loss to the company. This is the case when r^2=0. Higher utility loss occurs for r^2=5% than r^2=1%. This indicates that the company will suffer from greater utility loss if the bias against women increases.

We would multiply $\Delta U / \sigma$ by increasing function of rank to demonstrate the actual utility loss under different gender bias levels, assuming top levels of management have greater effect on corporation's bottom line. We would use two scenarios as follows:

Scenario I: `rank1<-c(1,2,3,4,5,6,7,8)`

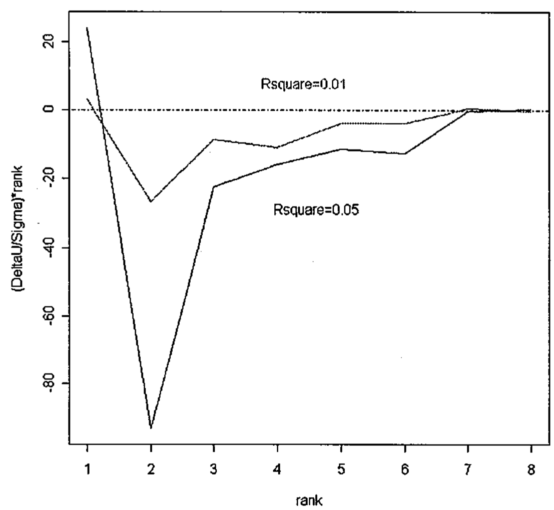
Say that CEO is 8 times as important as entry level;

Scenario II: `rank2<-c(1,1,2,2,3,3,4,4)`

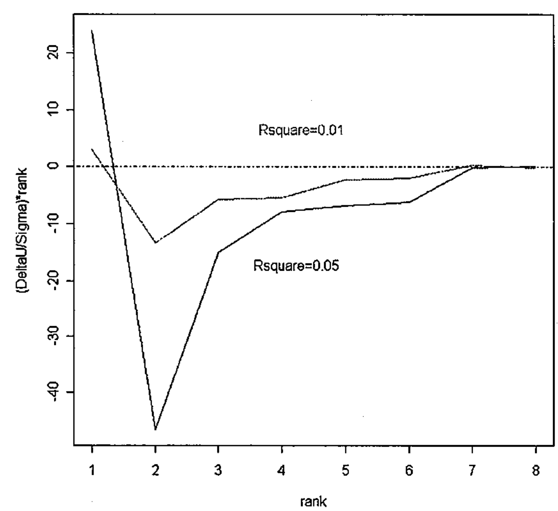
Say that CEO is 4 times as important as entry level.

The corresponding graphs under two scenarios are:

Scenario I



Scenario II



The curves far below the no-loss line further indicate that the company will suffer from greater utility loss if the bias against women increases. This might get the Human Resources department or the senior management team to think: what promotion policy or selection procedure should we employ in order to make the utility loss due to gender bias approximately zero?

SUMMARY

Decision making serves as the foundation on which utility theory rests (Fishburn, 1970, p. 1). As in decision under uncertainty, preferences between decision alternatives may be represented in terms of utilities for consequences and probabilities for consequences or for "states of the world" – which is the essential part in utility analysis in statistics and applied psychology. In complex situations, the computer simulation is a useful tool in assessing the utility gain or utility loss and thus helping decision making.

References:

Berger, James O. (1988), *Statistical Decision Theory, Foundations, Concepts, and Methods (second edition).* Springer-Verlag New York.

Brogden, H. E., & Taylor, E. K. (1950). The dollar criterion – Applying the cost accounting concept to criterion construction. *Personnel Psychology,* 3, 133-154.

Catalyst. (1998). *Census of women corporate officer's and top earner's.* New York: Catalyst.

Cox, D. R., Hinkley, D.V. (1974), *Theoretical Statistics*, Chapman and Hall.

Cronbach, L. J., & Gleser, G. C. (1957). *Psychological tests and personnel decisions* (2nd ed.). Urbana: University of Illinois Press.

Fishburn, Peter C. (1988), "Utility Theory", *Encyclopedia of Statistical Sciences,* Vol. 9, John Wiley & Sons

Fishburn, Peter C. (1970), *Utility Theory for Decision Making*, John Wiley & Sons.

Hull, J. C. (1928). *Aptitude testing*. Yonkers, NY: World Book.

Keeney, Ralph L., Raiffa, Howard (1976), *Decisions with Multiple Objectives: Preferences and Value Tradeoffs,* John Wiley & Sons.

Martell, R. F., & Robison-Cox, J. F. (2002), "From Bias to Exclusion, Why So Few Women at the Top? A Computer Simulation Approach", proposal submitted to NSF.

Schmidt, Frank L., Hunter John E., McKenzie Robert C., Muldrow Tressie W. (1979), "*Impact of Valid Selection Procedures on Work-Force Productivity*", *Journal of Applied Psychology* Vol. 64. No. 6, 609-626.

Schmidt, Frank L., Rauschenberger, J. (1986). "Utility Analysis for Practitioners: A Workshop". First Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, April 9.

Watson, Stephen R., Dennis M. Buede (1982), *Decision Synthesis, The Principles and Practice of Decision Analysis.* Cambridge University Press