質問項目作成法：SERVQUALを手がかりに

加藤淳一

要旨

今日，質問紙調査は卒業研究を始め様々なところで行われている。その裾野の広がりとともに，質問項目作成方法を学習しないままに質問項目を作成する状況が見られ始めている。大学で統計分析の方法を学んでも，質問項目作成法を学ぶ機会は提供できていない。

この状況を踏まえて，本研究ではマーケティングの学部学生を意識して質問項目作成手順を説明した。研究者の間では広く知られた質問項目としてParasuraman等のSERVQUALがある。そこでParasuraman et al.(1988) を中心に据えて，可能な限り学部学生が自身の力で質問項目を作成できるように方法を説明する。

この目的を達成するために，本研究は4章構成で論を進める。2章から5章までで，Parasuraman et al.(1988) の質問項目作成手順を11のステップにより説明する。このとき不足していると思われる説明を脚注で補足する。2章では，質問項目作成手順の11ステップの概略を整理する。

その後，3章では最初の質問項目をインタビューから作成する方法を説明する。4章では，2度に渡ってデータを収集し質問項目を洗練する方法を説明する。1度目は3章で作成した質問項目を洗練し，2度目は1度目で洗練した質問項目を一層洗練させる。

5章では，洗練された質問項目の信頼性と妥当性について評価方法を説明する。最後に6章で，本研究で示した質問項目作成の手順を整理し，実際に質問項目を作成しようとする学部学生へ向けての方向付けを示す。

キーワード：SERVQUAL　質問項目作成

1．導入

今日，質問紙調査は卒業研究を始め様々なところで行われている。その裾野の広がりとともに，質問項目作成方法を学習することなく質問紙を作成する。このような状況が見られ始めている。特に，大学で統計分析の方法を学んでも，質問項目作成法を学ぶ授業は提供できていない。

こうした状況を考慮したとき，マーケティングを担当する教員として，これまでのマーケティング研究の中から質問項目作成方法を説明できないかと考えるようになった。そのプロセスは，できるだけ具体的なステップとして示されるべきである。この要求に応える研究として，Parasuraman et al.(1988) に注目した。この研究は，サービスクオリティ測定を研究する研究者に広く知られた質問項目（SERVQUAL）の作成手順を説明している。このSERVQUALの作成方法からなら，

学生にとって学ぶことも多いと思われる。

だが，この研究はマーケティング研究者を対象としている。したがって，そのまま Parasuraman et al.(1988) を学部学生に手渡しただけでは，内容を十分理解して自身の卒業研究に役立てられると考えにくい。そこで本研究の目的は，次の１点である。

Parasuraman et al.(1988) を中心に据えて，可能な限り学部学生が自身の力で質問項目を作成できるように方法を説明する。

この目的を達成するために，本研究は４章構成で論を進める。２章から５章までで，不足していると思われる部分を脚注で補足しながらParasuraman et al.(1988) の質問項目作成手順を11のステップに従って説明する。２章では，質問項目作成手順の11ステップの概略を整理する。その後，３章では最初の質問項目をフォーカスグループインタビューから作成する方法を説明する。４章では，その質問項目を使ってデータを収集し尺度を洗練する方法を説明する。５章では，洗練された尺度の信頼性と妥当性について評価方法を説明する。最後に６章で，本研究で示した質問項目の作成手順を整理し，実際に質問項目を作成しようとする学部学生へ向けての方向付けを示す。

２．質問項目作成の11のステップの概略：Parasuraman et al.(1988)

本章では Parasuraman et al.(1988) で紹介された質問項目作成の11ステップの一つひとつを Parasuraman et al.(1988) に依拠しながら説明していく。まず，Parasuraman et al.(1988) に示された11ステップの概略図を図１に示す。

Parasuraman, et al. (1988) は，図１にあるような11のステップによりインタビューで収集した質問項目を以後のプロセスを通じて洗練させている。本研究は，これら11のステップを以下で大きく４つに分けて説明していく。ここでも図１の流れをこの４つに整理して俯瞰しておく。

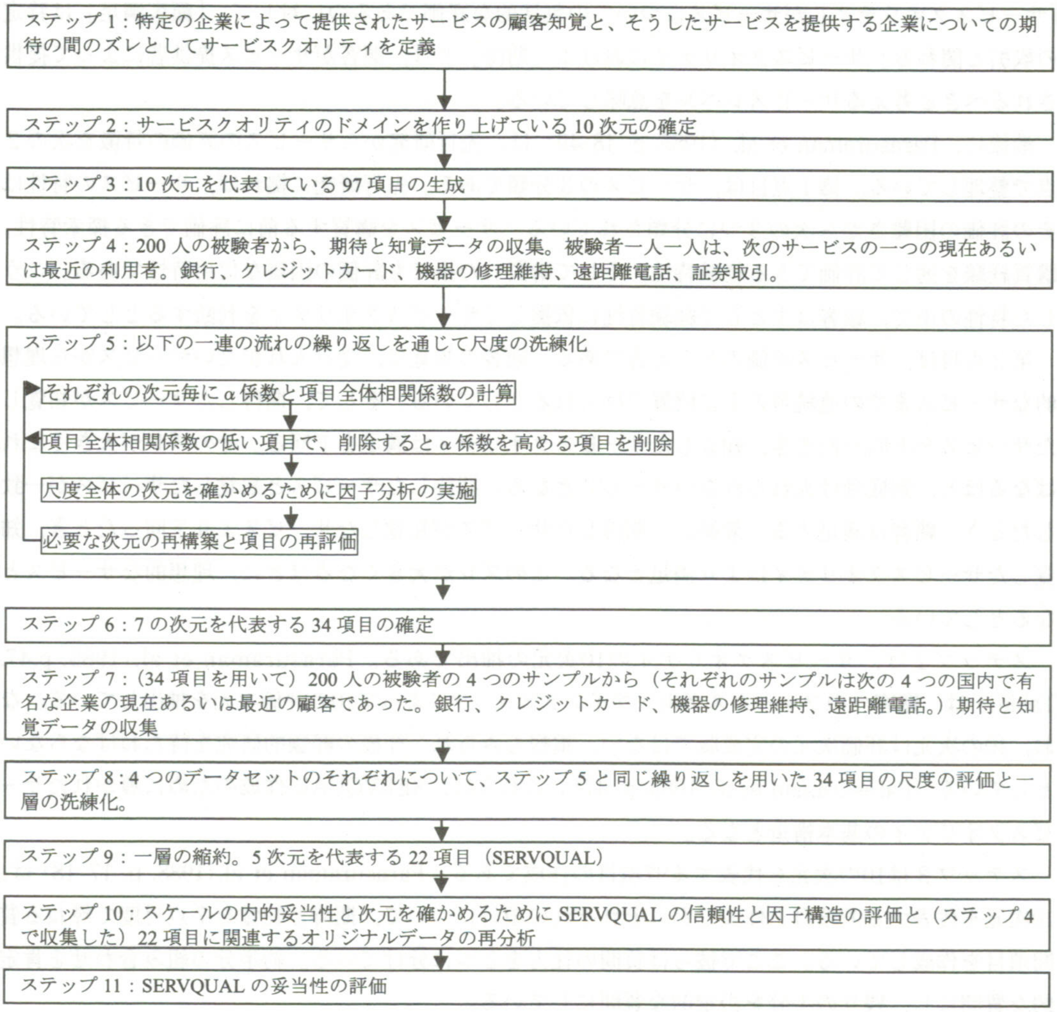
４つとは，ステップ１から３まで，ステップ４から６まで，ステップ７から９まで，そしてステップ10と11である。ステップ１から３までは，インタビューにより最初の質問項目を収集するところである。ここで収集された質問項目は97項目におよび，これを残りのステップで洗練させて質問項目を完成させている。

ステップ４から６までは，大きく２つのことをしている。まず，97項目の質問項目で質問紙調査を行い，データを収集している。次に，このデータを用いて，一貫した質問項目のみを残し，さらに重複した質問項目を削除している。こうして，質問項目が97項目から34項目へと洗練される。ステップ７から９までは，この34項目を用いて再度データを収集し，ステップ４から６までと同じ手順で34項目から22項目へと一層洗練させている。

ここまでで，２度に渡ってデータを収集して，質問項目を洗練させている。最後にステップ10と11では，この22項目の質問項目の信頼性と妥当性を確認している。以上の大きく４つにより，Parasuraman, et al. (1988) はサービスクオリティを測定する質問項目を完成させている。

以下で，より詳細に11のステップにしたがって，Parasuraman, et al.(1985, 1988) の質問項目の

図1　SERVQUAL 作成の11ステップ



出典：Parasuraman, et al. (1988), p.14, Figure 1.

作成手順を整理していく。なお，この11のステップの最初の部分は，Parasuraman et al.(1988）よりもむしろ Parasuraman et al.（1985）において詳細に語られている。したがって，ステップ1から3の途中まではParasuraman et al.（1985）の記述を中心にする。

3．97項目の作成

Parasuraman et al.（1988, p.15–18）は，インタビューから最初の質問項目として97項目を作成した。その97項目の作成までは，3つのステップでまとめられる。まず，これらのステップから説明する。

ステップ1は，サービスクオリティの定義である。Parasuraman et al.（1988, p. 15–17）は，用

語を次のように定義している。「知覚サービス」は，顧客の主観であって客観的な評価ではない。「サービスクオリティ」がサービスについての包括的な評価であるのに対して，「顧客満足」は特定の取引と関わる。サービスクオリティにおける「期待」とは，顧客がサービス提供者によって提供されるべきと考えるサービスレベルを意味している。

最後に，Parasuraman et al.（1985, p. 48–49）は，先行研究からサービスの評価の特徴を次の2点で整理している。第1点目は，サービスの3分類である。先行研究に拠れば，サービスの特性はその評価の困難さから次の3つに分類されている。サービスを購買する前に評価できる探索特性，購買経験を通じて評価できる経験特性，そして購買した後でも評価の困難な信用特性である。こうした特性の中で，顧客は主として経験特性に依拠してサービスクオリティを判断するとしている。

第2点目は，サービス評価のとらえ方である。顧客の知覚は，受け入れがたいサービスから理想的なサービスまでの連続帯の上に位置づけられるとしている。そして，期待したサービスが知覚したサービスを上回ったとき，知覚したサービスクオリティは満足を下回る。そのズレが大きくなればなるほど，到底受け入れられないサービスとなる。期待したサービスと知覚したサービスが一致したとき，顧客は満足する。最後に，期待したサービスが知覚したサービスより下回ったとき，知覚したサービスクオリティはより満足となる。このズレが大きくなるほどに，理想的なサービスとなるとしている。

ステップ2は，サービスクオリティの10次元の抽出である。Parasuraman et al.（1985, p.47, Table1）は，顧客へのフォーカスグループインタビュー[(1)]によって10の評価次元を抽出している。なお，10の次元は評価次元の完成版ではない。重複もみられ，今後の経験的研究を待たねばならないとしている（Parasuraman et al, 1985, p. 47)。これらは，SERVQUAL 作成のために導かれたサービスクオリティの基本構造となる。

ステップ3は10の次元を代表する97項目の作成である。Parasuraman et al.（1988, p. 17–18）は，10次元を代表する97項目を作成した。それぞれの項目について，期待と知覚を7点尺度で尋ねる質問項目を作成している。ここで彼らは質問の仕方を2つに分けている。約半分の組み合わせを肯定的な質問にし，残りの半分を否定的な質問にしている。

以上のようにして，Parasuraman et al.（1985, 1988）は最初の97項目を作成している。Parasuraman et al.（1988, p. 18–22）は，これらの項目についてデータを収集し尺度を洗練させている。データの収集と尺度の洗練は，つぎの3つのステップで行われている。この尺度の洗練によって，10次元97項目は7次元34項目まで絞り込まれている。

4．データの収集と尺度の洗練

（1）97項目でのデータ収集と尺度の洗練

ステップ4は，97項目についてデータの収集である。Parasuraman et al.（1988, p. 18）は，97項目を洗練するためのデータを次のようにして集めた。先行研究から代表的なサービス産業と見なせるものを5つ選んでいる。そして，そのどれかを過去3ヶ月以内に利用経験のある200名をサンプルとしている。サンプルは男女半々に分けられ，40名ずつ5つの異なるサービス産業に割り振られ

ている。彼らは，先の5つのうち3ヶ月以内に利用経験のあるサービス産業について先の97項目の質問紙に回答している。これをデータとしている。

ステップ5は，尺度の洗練である。Parasuraman et al.（1988, p. 19-20）は，10次元97項目の尺度を先のデータの分析により洗練した。まず，各次元内の一貫性を高める。Qスコア(2)を生データとしてクロンバックのα(3)が改善するまで，10次元の次元毎のクロンバックαの計算と10次元の次元毎の修正項目全体相関係数（the item's corrected item-all correlation）(4)の吟味を繰り返している。その結果，項目数は54項目へと減少している。

次に，次元間の重複を無くした。この課題は，54項目のQスコアの因子分析(5)によって検討されている。先験的に10の因子を仮定した主因子法（the principal axis factoring procedure）が用いられている。次元間の内的相関を許容するために，斜交回転（oblique rotation）にかけられ解釈しやすい因子負荷行列（factor loading matrix）を作成している。因子負荷行列をもとにサービスクオリティの次元を縮減している。ステップ5のここまでの手順（クロンバックのα，修正項目全体相関係数，そして因子分析）の反復により，当初の10次元から最終的に7つの次元で34項目へと絞られている。

ステップ6は，34項目が7の次元を代表していることを示した。Parasuraman et al.（1988, p.20）は，次のようにして34項目が7つの次元を代表していると示した。斜交回転に続いて，7次元の間の平均ペアワイズ相関係数（average pairwise correlation）(6)を求めている。平均ペアワイズ相関係数が低く，7つの異なる側面を持つことを示せたとしている。加えて34項目の結合信頼性（combined reliability）(7)は非常に高く，7つの次元の高い一貫性を示したとしている。

ここまでで7次元34項目まで絞られていた。Parasuraman et al.（1988, p. 22-24）は，7次元34項目を次の3つのステップによって更に洗練させている。次の洗練によって，7次元34項目は5次元22項目まで絞られる。そして22項目が，SERVQUALと呼ばれる。

（2）34項目でのデータの収集と尺度の洗練

ステップ7は，34項目についてデータの収集である。34項目を洗練させるためのデータが収集された。データは，全国的に知られた4つの企業それぞれについて先のステップ4と同じ条件で収集されている。同じ条件とは，各々の企業についてサンプルサイズ200名，25歳以上，その企業を過去3ヶ月以内に利用経験のある被験者からデータを取得したことを指している。ステップ4と同じく男女半々に分けられ，質問紙に答えてもらいデータとしている。

ステップ8は，34項目についてのさらなる洗練である。ここで，4つの企業についてのデータ分析により34項目の頑健性をテストしている。クロンバックのα，修正項目全体相関係数，そして因子分析を行っている。分析結果は，34項目をさらに洗練させなければならないことを示したとされる。そこで，修正項目全体相関係数の低かった項目を削除し，さらに因子分析結果を受けて2組の次元を統合している。結果として，次元は7次元から5次元へ減少している。

ステップ9は，5つの次元を代表する22項目（SERVQUAL）の作成である。5次元により，4つの企業についてのデータのクロンバックのαを再度計算し，斜交回転で因子分析を行っている。以後，ステップ5と同じ手順が繰り返されている。この分析の繰り返しにより，22項目5次元が作ら

れている。

以上で，５次元22項目からなる SERVQUAL が作成された。Parasuraman et al.（1988, p. 24-30）は，最後に２つのステップによってこの尺度の信頼性と妥当性の評価を行った。

５．尺度の信頼性と妥当性の評価

ステップ10は，SERVQUAL の信頼性，因子構造の評価と，ステップ４で集めたデータの再分析である。４つの企業すべてについて，クロンバックの α，因子負荷量のパターン，各項目の因子負荷量，そして５次元での平均ペアワイズ相関係数の全てで望ましい結果を得たとしている。この結果を受けて，ステップ４で集めたデータセットの中で22項目に該当する項目のみを取り出し再分析している。再分析により，次元の相関の低さと高い信頼性を再確認したとしている。こうして当初10次元97項目の尺度がSERVQUAL（５次元22項目）へと洗練されることで，次元の相関と信頼性の両者が改善されてきたことを確認している。

ステップ11は，SERVQUAL の妥当性の評価である。妥当性は，内容妥当性（content validity）(8)と収束妥当性（convergent validity）(9)によって評価されている。内容妥当性は，これまでの手順から満たされていると考えられたとしている。収束妥当性は，全体的評価とＱスコアの間で一元配置の分散分析とダンカン（Duncan）の方法(10)での多重比較を行っている。分析の結果，４つの企業をまたがって SERVQUAL の収束妥当性を強く支持したとしている。最後に，SERVQUAL によって測定された構成概念が，概念的に関連したその他の変数（他人に推薦したいかとトラブルはあったかの２つ）の測度と関連していたか否かを経験的に検討している。その結果，SERVQUAL の妥当性へさらなる支持を提供したと主張している。

６．結論

（１）本研究のまとめ

本研究では，学部学生が自身で理解できることを目指して質問項目の作成方法を説明してきた。本研究の最後に，質問項目の作成手順のポイントを整理しておく。フォーカスグループインタビューをもとにして，最初の質問項目を作成する（ステップ１－３）。これらの質問項目を使って，調査によりデータを得る（ステップ４）。

データから次元内の一貫性を高める。そのために，クロンバックの α や修正項目全体相関係数を繰り返し計算し，内的一貫性を高める項目のみを残す。次いで，因子分析により複数の次元に因子負荷量の高い項目を見直し，次元間の重複を無くしていく。これらを繰り返し求める（ステップ５）。更に，平均ペアワイズ相関係数により次元間の相関の低さを確認し，結合信頼性によって次元間の一貫性を確認する（ステップ６）。

こうして出来た尺度を用いてデータを集め直し，クロンバックの α，修正項目全体相関係数，そして因子分析を繰り返し求めて一層の洗練を行う（ステップ７－９）。この一層の洗練を経てできた尺度で，最初に収集したデータのうちで22項目に該当する項目のみを再び分析する（ステップ10）。この分析結果と当初の分析結果を比較することにより，これまでの過程によって尺度が洗練された

ことを確認する。最後に，妥当性を評価した（ステップ11）。こうして22項目の SERVQUAL が作られた。

以上で説明したように，11のステップを経て SERVQUAL は作成された。仮に質問項目を作成するならば，これらステップの一つひとつを理解して踏まなければならない。

（2）質問項目作成に向けて

本研究では，質問項目作成方法について SERVQUAL の作成手順に依拠して説明した。本研究で理解できた方法を卒業研究などに生かしてくれればと考えている。加えて，既に作られた質問項目を利用することの利点も強調しておく。質問項目が作れるようになったからといって，多くの研究者がむやみに質問項目を作ればよいというものではない。質問項目の作り方を理解したら，次の段階として既に作られてきた質問項目について知らなければならない。既に優れた質問項目が作られているならば，その質問項目を引き受けて追試をすることも重要である。

つまり，質問項目の作り方さえ知らずに質問紙調査を行っているようではいけない。だが，既に開発されている質問項目を調べもせずに，思いつきだけで様々な関連性のない質問項目を作るようではいけない。研究は常に積み上げられていくものであって，先人の貢献をきちんと読み評価しなければならない。

（かとう・じゅんいち　メディア社会学科）

参考文献

Cronbach L J. (1951), "Coefficient Alpha and the Internal Structure of Tests", *Psychometrika,* Vol. 16, No. 3, pp. 297–334.

三土修平（1997），『初歩からの多変量統計』，日本評論社。

永田靖・吉田道弘（1997），『統計的多重比較法の基礎』，サイエンティスト社。

Nunnally, J. C. and Ira H. Bernstein.（1994）, *Psychometric Theory 3rd Edition,* McGraw-Hill, Inc.

岡本安晴（2006），『計量心理学』，培風館。

Parasuraman, A., Valarie, A. Zeithaml and Leonard L. Berry. (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research", *Journal of Marketing,* Vol. 49 No. 4, pp. 41–50.

Parasuraman, A., Valarie, A. Zeithaml and Leonard L. Berry. (1988), "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality", *Journal of Retailing,* Vol. 67, No. 1, pp. 12–37.

Parasuraman, A., Valarie, A. Zeithaml and Leonard L. Berry. (1994), "Alternative Scales for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria", *Journal of Retailing,* Vol. 70, No. 3, pp. 201–230

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Vaughn, Sharon., Jeanne Shay Schumm, Jane M. Sinagub. (1996), *Focus Group Interviews in Education and*

Psychology, Sage Publications, Inc.（和訳：「グループ・インタビューの技法」，慶應義塾大学出版会）。

注

(1) Vaughn et al.（1996）によると，フォーカスグループインタビューの応用範囲の一つとして質問項目の開発も含まれており（Vaughn et al. 1996, p. 27），本研究の問題もフォーカスグループインタビューの目的としてかなっている。Vaughn et al.（1996, pp. 120-126, 和訳：pp. 154-163）は，フォーカスグループインタビューの手順を9つのステップに整理している。ステップ1は，全般的な目的の記述である。これは，4章（Vaughn et al, 1996, pp. 36-55，和訳：pp. 47-71）で詳細に論じられていた。全体的な目的の記述は，総合調査企画書の策定を意味する。

ステップ2は，司会者の選定である。これは，6章（Vaughn et al, 1996, pp. 74-94，和訳：pp. 97-123）で詳細に論じられる。Vaughn et al.（1996）は，有能な司会者の資質（Vaughn et al, 1996, p. 88, Table 6.3，和訳：p. 116）やリクルートするための質問（Vaughn et al, 1996, pp. 89, Table 6.4，和訳：p. 118）を表に整理している。ステップ3は，研究目的の精選すなわち総合調査企画書を一層具体的にすることである。これは，4章（Vaughn et al, 1996, pp. 36-55，和訳：pp. 47-71）で詳細に論じられていた。

ステップ4は，参加者の選定である。これは，5章（Vaughn et al, 1996, pp. 56-73，和訳：pp. 73-95）で詳細に論じられていた。フォーカスグループインタビューでの標本抽出の手続き（Vaughn et al, 1996, p. 58，和訳：p. 76），参加者の構成（Vaughn et al, 1996, p. 62，和訳：p. 80），募集の方法，インフォームドコンセントの文章例，さらに謝礼（Vaughn et al.（1996, pp. 64-73，和訳：pp. 84-95）にも言及している。ステップ5は，フォーカスグループインタビューのグループ数の決定である。これは，4章で詳細に論じられていた。グループ数の決定では，対象とするグループの数（Vaughn et al, 1996, p.49，和訳：p. 64），1グループの適切な人数（Vaughn et al, 1996, p. 50，和訳：p. 66），一度のインタビューの適切な時間（Vaughn et al, 1996, p. 50，和訳：p. 65），そして欠席者への備え（Vaughn et al, 1996, p. 51，和訳：p. 66）などに言及している。

ステップ6は，フォーカスグループインタビューを行う施設の手配である。これも，4章で詳細に論じられていた。施設を考える上で重要なポイントとして，Vaughn et al.（1996, pp.51-55，和訳：pp. 67-71）は次の4つをあげている。1つに，部屋の広さ。2つに，部屋が人を招く状態にあるか。3つに，部屋でどの程度記録用の機器を使えるか。そして最後に，参加者が喜んで参加してくれる場所にあるか。これは立地や駐車場の有無などを含んだ条件である。ステップ7は，インタビュー手引きの作成である。インタビュー手引きは，4章で詳細に論じられていた。導入から謝辞まで詳細な説明がなされている Vaughn et al,（1996, pp. 41-48，和訳：pp. 53-62）。

ステップ8は，フォーカスグループインタビューの実施である。これは6章で詳細に論じられていた。その説明は参加者が到着する前の会場の設置や確認に始まり，到着した参加者への対応，さらに討議を開始してからの討議の整理の仕方にまで及んでいる。ステップ9は，フォーカスグ

ループインタビューにおけるデータの分析である。これは，７章（Vaughn et al, 1996, pp. 95–117, 和訳：pp. 125–151）で詳細に論じられていた。代表的な考えの確認に始まり５段階で分析手順を整理している。このようにフォーカスグループインタビューを実施し質問項目の候補を出す。

(2)　Ｑスコアは，Q＝P-Eによって求められる。ここでＰは知覚の観測値，Ｅは期待の観測値である。

(3)　クロンバックのαは，信頼性とくに内的一貫性（internal consistency）を測定する係数である。ここで信頼性とは，テストの時間的な安定性やテスト項目間の一貫性を指している。内的一貫性とは，テストを構成している項目の中に無関係な項目の混在していないことを意味している。クロンバックのαは次式で与えられ，値が高いほど内的に一貫している。

$$\alpha = \frac{k}{k-1}\left(1-\frac{\sum_{j=1}^{k} s_j^{\,2}}{s_X^{\,2}}\right)$$

上の式の表記を説明する。全被験者ｎ人に質問紙調査を行い，その質問紙のうちｋ項目の一貫性を調べているとする。たとえば，ある質問紙が３つの次元で構成されており，次元１が４項目，次元２が５項目，次元３が４項目であるとする。この次元２の内的一貫性について考察するなら，k＝5で計算する。ではクロンバックのαの表記を説明する。まず，$s_j^{\,2}$とは，全被験者ｎ人について，質問項目ｊの得点の不偏分散である。式では$s_j^{\,2}=\frac{\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2}{n-1}$のようになる。$s^2x$は，全被験者ｎ人について，被験者ｉの全質問項目のうちｋ個の質問項目の得点の合計点 X_iの不偏分散である。式では$s_X^{\,2}=\frac{\sum_{i=1}^{n}(X_i-\bar{X}_i)^2}{n-1}$のようになる。

ここからクロンバックのαの考え方を岡本（2006　補足5–1）と Nunnally and Bernstein（1994）に依拠しつつ著者なりに説明する。基本となるのは次の分散比である。この比が質問項目の一貫性を示す指標であり，クロンバックのαである。この分散比は，全ての被験者の全ての項目の分散で，全ての被験者の真の値の分散を割った分散比$\rho=\frac{\sigma^2{}_{T_i}}{\sigma^2{}_{X_i}}$である。観測値Ｘを真の値Ｔに誤差ｅを加えたものと仮定する。T^*は，被験者ｉの各々の質問項目の真の値から，被験者ｉのｋ個の質問項目の真の値の平均点を引いた差と定義し，どの質問項目の場合も同じT^*と仮定する。つまり，$T^*=T_{i1}-\mu_{Ty}=\cdots T_{ik}-\mu_{Ty}$とする。最後に，誤差項ｅはランダムでＴと無相関と考える。したがって，$E[(T-\mu)(T-\mu)+e(T-\mu)]$ の$E[e(T-\mu)]$は，次のようにｅとＴの共分散$E[\{e-E(e)\}\{T-E(T)\}]=E[(e-0)(T-\mu)]=E[e(T-\mu)]=0$なのでゼロとなる。

まず，分散比の分子について検討する。この分散比$\rho=\frac{\sigma^2{}_{T_i}}{\sigma^2{}_{X_i}}$からクロンバックの$\alpha$の式を導出するに当たり，とりわけ分子の真の値の分散という直接観察できないものをいかにして計算できるようにするかがポイントとなる。直接観測できない真の値の分散を次のように推定する。

$$\begin{aligned}
\sigma^2{}_{T_i} &= E[\{(T_{i1}+\ldots+T_{ik})-(\mu_{T_{ij}}+\ldots+\mu_{T_{ij}})\}^2] \\
&= E[\{\sum_{j=1}^{k}(T_{ij}-\mu_{T_{ij}})\}^2] \\
&= E[(kT^*)^2] \\
&= k^2E[(T^*)^2] \\
&= k^2\sigma^2{}_{T^*} \qquad (1)
\end{aligned}$$

式展開の最後の部分は，被験者ｉの全ての質問項目の真の値から，それぞれの平均値を引いた値T^*の分散を$E[\{T^*-E[T^*]\}^2]=E[(T^*)^2]=\sigma^2{}_{T^*}$として求めている。なお，ここで$E[T^*]=0$であることに注意を要する。なぜならば$T^*$が，被験者ｉの各々の質問項目の真の値から，被験者ｉのｋ個の質問項目の真の値の平均点を引いた差である。したがって，平均からの差を平均（あるいは，合計）すればゼロになる。

端的には平均点50点で学生数２名のクラスを考える。１名は100点で，もう１名は０点とする。このとき『(100－50)＋(０－50)』÷２が平均からの差の平均となっている。すると，50－50=0となる。したがって，$E[T^*]=0$となる。この式を言葉にすると，真の得点の分散は，真の得点の推定値の分散をｋ^{2}倍した値と等しい。

さらに，x_{ij}とx_{il}の共分散をもとめると次の式変形が成り立つ。

$$
\begin{aligned}
&E[\{x_{ij}-E(x_{ij})\}\{x_{il}-E(x_{il})\}]\\
&=E[(x_{ij}-\mu_{T_{ij}})(x_{il}-\mu_{T_{il}})]\\
&=E[(T_{ij}+e_{ij}-\mu_{T_{ij}})(T_{il}+e_{il}-\mu_{T_{il}})]\\
&=E[(T_{ij}-\mu_{T_{ij}})(T_{il}-\mu_{T_{il}})]\\
&=E[(T^*)^2]\\
&=\sigma^2{}_{T^*} \qquad (2)
\end{aligned}
$$

この式変形で，誤差e_{ij}の平均をゼロとすると，$E[x_{ij}]=E[T_{ij}+e_{ij}]=E[T_{ij}]+E[e_{ij}]=E[T_{ij}]=\mu_{ij}$となる。この式から観測できないT*の分散$\sigma^2{}_{T^*}$の推定値として$x_{ij}$と$x_{il}$の共分散$S_{jl}$を用いる。$S_{jl}$は，１つの質問項目ｊと１つの質問項目ｌの共分散である。このままでは，ｊとｌは任意の１つの質問項目を指しており計算のときに困る。そこで，ｋ個の全ての質問項目について共分散を求め，その平均値をもって観測できない$\sigma^2{}_{T^*}$の推定値とする。つまり，$\sigma^2{}_{T^*}=\frac{1}{k(k-1)}\sum_{j\neq l}^{k}s_{jl}$となる。ここでｊとｌが別々の質問項目を指していると想定しているので，平均するときｋ^{2}ではなくｋ^2－ｋ＝ｋ（ｋ－１）で割る。このｋはｋ×ｋの分散共分散行列における対角成分にある分散の個数である。分散共分散行列の要素の個数はｋ^{2}個で，そのうち分散のｋ個を引いて，共分散の個数はｋ^2－ｋ＝ｋ（ｋ－１）個となる。以上から，観測不能な真の値の分散を観測可能な共分散の平均で推定すると次式のようになる。

$$\sigma^2{}_{T_i}=k^2\times\sigma^2{}_{T^*}=k^2\times\frac{1}{k(k-1)}\sum_{j\neq l}^{k}s_{jl} \qquad (3)$$

以上から，分散比$\rho=\frac{\sigma^2{}_{T_i}}{\sigma^2{}_{X_i}}$の分子は式(3)を代入する。

次に，分散比の分母について考える。X_i（被験者ｉのｋ個の質問項目の得点の合計）の全被験者ｎ人についての分散をs^2x（$=\sigma^2x_i$）とすると，次の式変形により共分散で表現できる。

$$
\begin{aligned}
&\sigma^2{}_{X_i}=s^2x=E[\{X_i-E(X_i)\}^2]\\
&=E[(X_i-\overline{X}_i)^2]\\
&=E[\{(x_{i1}+\ldots+x_{ik})-(\mu_{T_{i1}}+\ldots+\mu_{T_{ik}})\}^2]\\
&=\sum_{j,l}^{k}E[(x_{ij}-\mu_{T_{ij}})(x_{il}-\mu_{T_{il}})] \qquad (4)
\end{aligned}
$$

この式変形の$E[\{(x_{i1}+\ldots+x_{ik})-(\mu_{T_{i1}}+\ldots+\mu_{T_{ik}})\}^2]=\sum_{j,l}^{k}E[(x_{ij}-\mu_{T_{ij}})(x_{il}-\mu_{T_{il}})]$の変形は分散で表現された式を式(2)で示されているように質問項目ｊと質問項目ｌの平均からの差が等しいことを利用して共

分散に変形した。さらに，式(4)の $E(X_i)$ は，k 個全ての得点の合計（X_i）を被験者の人数 n 人分計算し，その n 人分の合計得点（$\sum_{i=1}^{n} X_i$）の平均点 $E(X_i)=\frac{\sum_{i=1}^{n} X_i}{n}=\frac{\sum_{i=1}^{n}\sum_{j=1}^{k} x_{ij}}{n}$ となる。したがって，式(4)より，X_i の分散 s^2x は質問項目 j の得点 x_{ij} と質問項目 1 の得点の全被験者 n の共分散で表現できる。そして，を全被験者n人について，質問項目 j と質問項目 1 の共分散とする。すると，X_i の分散 s^2x は次のようになる。

$$s^2 x=\sum_{j,l}^{k} s_{jl} \quad (5)$$

以上から，分散比 $\rho=\frac{\sigma^2 T_i}{\sigma^2 x_i}$ は，つぎのようになる。まず，分母 $\sigma^2 x_i$ は X_i の分散を意味している。ここまでの表記にしたがえば式(5)より，分母 $\sigma^2 x_i$ は $s^2 x=\sum_{j,l}^{k} s_{jl}$ となる。次に，分子 $\sigma^2 T_i$ は式(3)から $\sigma^2 T_i = k^2 \times \frac{1}{k(k-1)}\sum_{j\neq l}^{k} s_{jl}$ となる。これらを代入して，計算すると

$$\begin{aligned}\rho&=\frac{\sigma^2 T_i}{\sigma^2 x_i}=\frac{k^2\frac{1}{k(k-1)}\sum_{j\neq l}^{k} s_{jl}}{\sum_{j,l}^{k} s_{jl}}\\&=\frac{\frac{k}{(k-1)}\sum_{j\neq l}^{k} s_{jl}}{\sum_{j,l}^{k} s_{jl}}=\frac{k}{k-1}\times\frac{\sum_{j,l}^{k} s_{jl}-\sum_{j=1}^{k} s_{jj}}{\sum_{j,l}^{k} s_{jl}}\\&=\frac{k}{k-1}\times\left(1-\frac{\sum_{j=1}^{k} s_{jj}}{\sum_{j,l}^{k} s_{jl}}\right)=\frac{k}{k-1}\left(1-\frac{\sum_{j=1}^{k} s^2{}_j}{s^2 x}\right)\end{aligned} \quad (6)$$

式(6)の式変形の中で，$\sum_{j\neq l}^{k} s_{jl}$ が $\sum_{j,l}^{k} s_{jl}-\sum_{j=1}^{k} s_{jj}$ となる理由を説明する。共分散の和（$\sum_{j\neq l}^{k} s_{jl}$）で j ＝ 1 を認めると，分散共分散行列の全要素の和となり分散の分だけ余計となる。そこで，分散の分だけ引いて，共分散の和（$\sum_{j\neq l}^{k} s_{jl}$）が分散共分散行列の全成分の和から分散の成分の和を引いた（$\sum_{j,l}^{k} s_{jl}-\sum_{j=1}^{k} s_{jj}$）差となる。なお，$\sum_{j,l}^{k} s_{jl}-\sum_{j=1}^{k} s_{jj}$ の $\sum_{j=1}^{k} s_{jj}$ は全被験者 n 人について質問項目 j の得点の分散を S_j^2 と表記した。

よって，クロンバックの α は，次式で与えられる。

$$\alpha=\frac{k}{k-1}\left(1-\frac{\sum_{j=1}^{k} s_j{}^2}{s_x{}^2}\right)$$

ここでは，Cronbach (1951) に依拠して，既に示した α と他の関連指標について手短に整理する。あるテストの信頼性は，２度同じテストを繰り返して結果を比較すればよい。これを再テスト法という。だが，１度でも大変なテストを２度繰り返すのは一層難しい。そこで，折半法（split-half approach）が提案された。折半法とは，１つのテストを例えば偶数と奇数の項目のように２つに分けて２度テストを繰り返す代わりとする手法である。だが，再テスト法と折半法とは，同じものを測っていないと批判された。

テストの正確性は，安定性（stability）と同等性（equivalence）に分けられる。安定性とは，

時間がたっても同じように測定されることを指す。対して，同等性とは質問項目間の一貫性を指す。再テスト法は，安定性である。だが折半法は，同等性を測っている。

ここで考察の範囲を同等性のみに集中する。Kuder Richardson は1回のテストから得られたデータをもちいて同等性の係数を導いた。その公式が Kuder-Richardson 20 Formula である。KR20 は，次の式で与えられる。

$$r_{tt(KR20)} = \frac{n}{n-1}\left(1 - \frac{\sum_i p_i q_i}{\sigma^2_t}\right), \qquad (i = 1,2,\ldots,n).$$

i は項目を表し，p_i は項目 i について1をとった割合で，q_i は0をとった割合である。つまり KR20 は，1か0かの2択のケースにしか使えない。これに対して先に示した α は，より一般的な式として提案された。

ただ，これまで KR20 が折半法と同じものを測定できているのか示されていなかった。そこで Cronbach (1951) は，α が折半法の平均と一致することを発見した。こうして α が折半法で求めていた同等性を測定する指標として用いられることを示した。

(4) 修正項目全体相関係数は，複数の項目の中に無関係な項目が混在していると疑われるとき，どの項目が無関係なのかを相関係数を計算することで見つけ出すための係数である。

修正項目全体相関係数は，次のように計算される。j は項目を表し，i は被験者を表す。T_i は被験者 i の全ての項目の合計得点（$T_i = \sum_j T_{ij}$），a_i は被験者 i の項目 j の得点を表す。Cov（$T_i - a_i, a_i$）は全被験者の $T_i - a_i$（全ての項目の合計得点から項目 j の得点を引いた得点）と全被験者の a_i（項目 j の得点）の共分散を表す。S_{Ti-ai} は全被験者の $T_i - a_i$（全ての項目の合計得点から項目 j の得点を引いた得点）の標準偏差，S_{ai} は全被験者の a_i（項目 j の得点）の標準偏差を表す。

$$CITC = \frac{Cov(T_i - a_i, a_i)}{S_{T_i - a_i} S_{a_i}}$$

この係数を求めるとき，逆転項目が含まれていないように注意しなければならない。この相関係数が低いと，無関係な項目が混在していると疑われる。無関係な項目が同じ次元内に入っているならば，当該項目を除くことで次元内の一貫性はより高まると考えられる。一貫性が高まったか否かは，修正項目全体相関の低い項目を取り除いた後の新しい項目でクロンバックの α を計算し直せば分かる。

(5) 因子分析は，潜在的な因子を想定し観測値をある個数の因子でとらえる分析手法である。本研究の目的は，マーケティングの授業で扱えきれない質問項目作成に関連した知識を説明することである。因子分析は，マーケティングで必要とされる統計学の知識であり一般に授業でも扱われている。また，多数の入門書で学習できるようになっている。したがって，本研究ではその手法の詳細は取り上げない。たとえば三土（1997）を参照していただきたい。さらに R Development Core Team（2009）のように，フリーの統計ソフトも公開されている。

(6) 平均ペアワイズ相関係数を調べたが，定義・計算式を見つけられなかった。Parasuraman, et al.（1988）の前後関係から次のように理解できた。例として3つの次元を想定する。次元1，次元2，そして次元3と呼ぶ。次元1は4つの質問項目で構成されている。次元2は，3つの質問

項目で構成されている。最後に，次元３は４つの質問項目で構成されている。したがってこの質問紙は，３つの次元11の質問項目で構成されている。

さて，Parasuraman, et al.(1988) は「質問紙の平均ペアワイズ相関係数が低いことから，次元が異なる側面をとらえている」と結論づけていた。つまり，平均ペアワイズ相関係数は，次元間での比較をするための相関係数となる。比較は，３つの組み合わせで考えられる。次元１と次元２，次元１と次元３，次元２と次元３となる。

各々の次元を構成している質問項目について，多数の被験者の調査結果をもとめる。この調査結果を被験者ごとに，次元ごとで平均する。たとえば，被験者Ａさんの次元１の４つの質問項目の得点の平均値，次元２の３つの質問項目得点の平均値，次元３の４つの質問項目得点の平均値を計算する。被験者人数分だけ，次元ごとに平均値を計算し，これらの平均値間の相関係数を計算する。表に整理すると，次のようになる。

被験者	次元１	次元２	次元３
A	Ａさんの次元１の４つの質問項目の平均値	Ａさんの次元２の３つの質問項目の平均値	Ａさんの次元３の４つの質問項目の平均値
B	Ｂさんの次元１の４つの質問項目の平均値	Ｂさんの次元２の３つの質問項目の平均値	Ｂさんの次元３の４つの質問項目の平均値
C	Ｃさんの次元１の４つの質問項目の平均値	Ｃさんの次元２の３つの質問項目の平均値	Ｃさんの次元３の４つの質問項目の平均値
D	Ｄさんの次元１の４つの質問項目の平均値	Ｄさんの次元２の３つの質問項目の平均値	Ｄさんの次元３の４つの質問項目の平均値
E	Ｅさんの次元１の４つの質問項目の平均値	Ｅさんの次元２の３つの質問項目の平均値	Ｅさんの次元３の４つの質問項目の平均値
⋮	⋮	⋮	⋮

複数の被験者ごとに計算された平均値をデータとして
次元１と次元２の相関係数をもとめる。

そして，この相関係数を相関行列にして考える。相関行列を例示してみると，次のようになる。この相関行列の対角成分以外の成分の平均が，平均ペアワイズ相関係数となる。このように考えると，この相関係数が低い場合，次元１から３が相互に次元は異なる側面をとらえていると結論づけられる。

	次元１	次元２	次元３
次元１		次元２と１の相関	次元３と１の相関
次元２	次元１と２の相関		次元３と２の相関
次元３	次元１と３の相関	次元２と３の相関	

(7) Nunnally and Bernstein（1994, p. 266）は，結合信頼性についてReliability of Liner Combinationsという節で言及している。結合信頼性も，クロンバックのαで出発点にした分散比から始める。クロンバックのαの説明との相違は，分母・分子ともに線形結合した$X=x_1+x_2+\cdots+x_5$と$T=t_1+t_2+\cdots+t_5$のような変数になるところである。先に注で説明したクロンバックのαはParasuraman et al.（1988, p. 21, Table 1）において，5つの次元ごとに計算され，合計で5つのクロンバックのαが計算される。これらのクロンバックのαは，尺度全体ではなく，次元内が一貫しているか否かを調べている。これに対して，ここで結合信頼性は例えば5次元の線形結合として尺度全体の信頼性を求める。したがって，各次元のクロンバックのαが求められた後に，この結合信頼性が計算される。

ここでは一般に結合信頼性の考え方を説明する。結合信頼性も，$\alpha=\dfrac{\text{真の値の分散}}{\text{観測データの分散}}$から始まる。この分散比において分母・分子ともに線形結合した変数なので，

$\alpha=\dfrac{\text{真の値の線形結合}(T)\text{分散}}{\text{観測データの線形結合}(X)\text{分散}}$となる。分母（$\sigma^2x$）は，Nunnally and Bernstein（1994, p. 161）から観測データx_1，x_2，…，x_5の分散共分散行列の全要素の和と一致する。分子（σ^2T）は真の値t_1，t_2，…，t_5の分散共分散行列の全要素の和と一致する。さらにクロンバックのα説明中に出てきた式(2)から，真の値の共分散と観測データの共分散が一致する。したがって，真の値の分散共分散行列の全要素の和（σ^2T）は，観測データの分散共分散行列の和（σ^2x）から観測データの分散共分散行列の分散部分の和（$\sum\sigma^2{}_{x_i}$）を引き，真の値の分散共分散行列の分散部分の和（$\sum\sigma^2{}_{t_i}$）を加えることで表せる。式にすると，次のようになる。

まず次元ごとの信頼係数を観測データで表す。すると，次元ごとの信頼性係数は$\alpha_i=\dfrac{\sigma^2{}_{t_i}}{\sigma^2{}_{x_i}}$より$\sigma^2_{Ti}=\alpha_i\sigma^2x_i$となり，次元ごとの真の値の分散は個々の信頼係数と観測データの分散との積で表せる。これより，分子は$\sigma^2{}_T=\sigma^2X-\sum\sigma^2{}_{x_i}+\sum\sigma^2{}_{t_i}=\sigma^2X-\sum\sigma^2{}_{x_i}+\sum\alpha_i\sigma^2{}_{x_i}$となる。

以上から，結合信頼性は次式で求まる。

$$
\begin{aligned}
\alpha &= \frac{\text{真の値の線形結合}\quad (T)\text{分散}}{\text{観測データの線形結合}\quad (X)\text{分散}} \\
&= \frac{\sigma^2{}_T}{\sigma^2X} \\
&= \frac{\sigma^2X-\sum\sigma^2{}_{x_i}+\sum\alpha_i\sigma^2{}_{x_i}}{\sigma^2X} \\
&= 1-\frac{\sum\sigma^2{}_{x_i}-\sum\alpha_i\sigma^2{}_{x_i}}{\sigma^2X}
\end{aligned}
$$

この式から各次元間を跨がった尺度全体の信頼性は次のように読み替えて計算できる。各次元の分散（σ^2x_i）（特に言及は見当たらないが，各次元を構成している質問項目の線形結合の分散），各次元のクロンバックのα（α_i），そして各次元の分散共分散行列の全要素の和（σ^2x_i）によって，結合信頼性が求められる。

(8)「内容的妥当性（content validity）は測ろうとするものについての概念的規定を明らかにして，その規定に基づいて当該の概念が測定されていると判断できれば内容的妥当性があるということ

になる（岡本，2006，p. 169)」。

(9)「同じ特性（構成概念）を測っていると考えられる尺度間の相関係数は高くなければならないというのが収束的妥当性である（岡本，2006，p. 171)」。

(10) Duncanの方法は，用いてはいけないとされている。例えば永田・吉田（1997，p. 31）に詳しい。

Procedure to Create Items of Questionnaire: Lessons from SERVQUAL

Junichi Kato

Nowadays, many undergraduate students use questionnaire in their graduate these. As the survey methods get popular, some students create the questionnaire without learning how to compile the items. We cannot teach how to make items in classes of undergraduate level, even if we can teach students the statistical methods in classes.

Under this circumstance, this paper explains how to make questionnaire for undergraduate students who major in marketing. SERVQUAL scale (Parasuraman et al., 1988) is the well-known questionnaire to most marketing researchers, but it is difficult for undergraduate students. So this paper supports them to understand how to create questionnaire by their own power.

To achieve this goal, this paper is constituted of four chapters. In chapter 2 to 5, we explain the procedure to make items on the basis of Parasuraman et al. (1988). We give some supplementary explanation in footnotes for easy understanding.

Keyword：SERVQUAL, Questionnaire