

O uso de questionários em trabalhos científicos

1. Introdução

Um questionário é tão somente um conjunto de questões, feito para gerar os dados necessários para se verificar se os objetivos de um projeto foram atingidos. Mas, construir questionários não é uma tarefa fácil, e aplicar tempo e esforço no planejamento do questionário é um requisito essencial para se atingir os resultados esperados. Infelizmente não existe uma metodologia padrão para o desenvolvimento de questionários, porém existem recomendações de diversos autores com relação a essa importante tarefa no processo de pesquisa científica.

Este trabalho tem o objetivo de apresentar as etapas envolvidas no desenvolvimento de um questionário e está estruturado da seguinte forma: a seção 2 versa sobre a metodologia para a construção de um questionário; a seção 3 apresenta a importância de se definir uma amostra adequada de respondentes e como definir o tamanho da amostra; a seção 4 descreve uma visão geral de como analisar os dados coletados pelos questionários; e finalmente a seção 5 apresenta as conclusões.

2. Metodologia para a construção de um questionário

A construção de um questionário, segundo Aaker et al. (2001), é considerada uma “arte imperfeita”, pois não existem procedimentos exatos que garantam que seus objetivos de medição sejam alcançados com boa qualidade. Ainda segundo o autor, fatores como bom senso e experiência do pesquisador podem evitar vários tipos de erros em questionários, como por exemplo, as questões ambíguas, potencialmente prejudiciais, dada sua influência na amplitude de erros. No entanto, existe uma seqüência de etapas lógicas que o pesquisador deve seguir para desenvolver um questionário:

- 1 – Planejar o que vai ser mensurado
- 2 – Formular as perguntas para obter as informações necessárias.
- 3 – Definir o texto e a ordem das perguntas e o aspecto visual do questionário.
- 4 – Testar o questionário, utilizando uma pequena amostra, em relação a omissões e ambigüidade.
- 5 – Caso necessário, corrigir o problema e fazer novo pré-teste.

Baseada em Aaker et al. (2001), a Figura 1 ilustra os passos para elaboração de um questionário:

Etapa	Passos
Planejar o que vai ser Mensurado	Evidenciar os objetivos da pesquisa
	Definir o assunto da pesquisa em seu questionário
	Obter informações adicionais sobre o assunto da pesquisa a partir de fontes de dados secundários e pesquisa exploratória
	Determinar o que vai ser perguntado sobre o assunto da pesquisa
Dar Forma ao Questionário	Para cada assunto, determinar o conteúdo de cada pergunta
	Decidir sobre o formato de cada pergunta
Texto das Perguntas	Determinar como as questões serão redigidas
	Avaliar cada uma das questões em termos de sua facilidade de compreensão, conhecimentos e habilidades exigidos, e disposição dos respondentes.
Decisões sobre Seqüenciamento e Aparência	Dispor as questões em uma ordem adequada
	Agrupar todas as questões de cada sub-tópico para obter um único questionário
Pré-Teste e Correção de Problemas	Ler o questionário inteiro para verificar se faz sentido, e se consegue mensurar, o que está previsto para ser mensurado
	Verificar possíveis erros no questionário
	Fazer o pré-teste no questionário
	Corrigir o problema

Figura 1 – Passos para a elaboração de um questionário.

2.1 Decisões sobre o conteúdo das perguntas

Com relação ao conteúdo das perguntas, pode-se tentar verificar fatos, crenças quanto a fatos, crenças quanto a sentimentos, descoberta de padrões de ação e de comportamento presente ou passado. Também é necessário que o pesquisador faça algumas reflexões, do tipo: a pergunta é realmente necessária? Qual a sua utilidade? Estas perguntas podem se desdobrar nas seguintes questões:

- O assunto exige uma pergunta separada, ou pode ser incluído em outras perguntas?
- Existem outras perguntas que já incluem adequadamente este ponto?
- A pergunta é desnecessariamente minuciosa e específica?
- Várias perguntas são necessárias sobre o assunto desta pergunta ou uma é o suficiente?
- Deve-se evitar o uso de abreviação, e não se deve tratar dois assuntos complexos em uma mesma pergunta.
- Todos os aspectos importantes sobre este tópico serão obtidos da forma como foi elaborada a pergunta?
- Em perguntas de opinião, interessa saber os graus de favorabilidade/desfavorabilidade, ou basta saber se é a favor ou contra?
- As pessoas têm a informação necessária para responder a pergunta? Não basta, porém que se esteja abordando a pessoa certa, é preciso saber se ela é capaz de se lembrar da informação. Nossa habilidade para nos lembrarmos dos eventos é influenciada pela importância do próprio evento para cada um, do tempo passado desde que ele ocorreu e da presença de estímulos que nos ajudem a recordar.
- Os respondentes estarão dispostos a dar a informação? A predisposição em responder a uma determinada questão parece ser função do tempo e trabalho envolvidos na elaboração da resposta, de sua habilidade em articular a resposta, e da sensibilidade do assunto tratado.
- Que objeções alguém poderia ter para responder esta pergunta?
- O tema abordado é muito íntimo, perturbador ou expõe socialmente as pessoas, de forma a causar resistências e respostas falsas?
- O tema é embaraçoso para o respondente por colocar em perigo seu prestígio caso seja contrário a idéias socialmente aceitas?

Deve-se tomar o cuidado de não se usar perguntas muito específicas, quando, na verdade, a pesquisa for de caráter geral. Por exemplo, perguntar quantas vezes uma pessoa foi ao supermercado em determinado mês, pode resultar em uma resposta menos precisa do que se fosse perguntado a respeito do seu comportamento usual ou médio durante os meses anteriores.

- O conteúdo da pergunta não estará enviesado ou carregado em determinada direção?

Esta pergunta desdobra-se nas seguintes questões:

- A pergunta é, devidamente, neutra, a fim de não influenciar nas respostas?
- Pessoas com opiniões contrárias sobre o assunto não a considerarão tendenciosa?
- A pergunta contém opiniões ou julgamentos relacionados ao assunto?

2.2 Decisões sobre o formato das respostas

A escolha do formato das respostas mais adequado deve levar em conta as vantagens e desvantagens de cada tipo para o objetivo da pesquisa.

As questões podem ser:

- abertas – onde os respondentes ficam livres para responderem com suas próprias palavras, sem se limitarem a escolha entre um rol de alternativas.
- múltipla escolha - onde os respondentes optarão por uma das alternativas, ou por determinado número permitido de opções.
- dicotômicas - São as que apresentam apenas duas opções de respostas, de caráter bipolar, do tipo: sim/não; concordo/não concordo; gosto/não gosto. Por vezes, uma terceira alternativa é oferecida, indicando desconhecimento ou falta de opinião sobre o assunto.

A figura 2 descreve segundo Mattar (1994), as principais vantagens e desvantagens de cada formato das respostas.

Tipo de Questões	Vantagens	Desvantagens
Abertas	<ul style="list-style-type: none"> ➤ Estimulam a cooperação; ➤ Permitem avaliar melhor as atitudes para análise das questões estruturadas; ➤ São muito úteis como primeira questão de um determinado tema porque deixam o respondente mais à vontade para a entrevista a ser feita; ➤ Cobrem pontos além das questões fechadas; ➤ Têm menor poder de influência nos respondentes do que as perguntas com alternativas previamente estabelecidas; ➤ Exigem menor tempo de elaboração; ➤ Proporcionam comentários, explicações e esclarecimentos significativos para se interpretar e analisar as perguntas com respostas fechadas; ➤ Evita-se o perigo existente no caso das questões fechadas, do pesquisador deixar de relacionar alguma alternativa significativa no rol de opções. 	<ul style="list-style-type: none"> ➤ Dão margem à parcialidade do entrevistador na compilação das respostas, já que não há um padrão claro de respostas possíveis. Assim, é difícil a codificação das respostas e sua conseqüente compilação; ➤ Há grande dificuldade para codificação e possibilidade de interpretação subjetiva de cada decodificador; ➤ Quando aplicadas em forma de entrevistas, podem levar potencialmente a grandes vieses dos entrevistadores; ➤ Quando feitas através de questionários auto-preenchidos, esbarram com as dificuldades de redação da maioria das pessoas, e mesmo com a "preguiça" de escrever. ➤ São menos objetivas, já que o respondente pode divagar e até mesmo fugir do assunto; ➤ São mais onerosas e mais demoradas para serem analisadas que os outros tipos de questões.
Múltipla Escolha	<ul style="list-style-type: none"> ➤ Facilidade de aplicação, processo e análise; ➤ Facilidade e rapidez no ato de responder; ➤ Apresentam pouca possibilidade de erros; ➤ Diferentemente das dicotômicas, trabalham com diversas alternativas. 	<ul style="list-style-type: none"> ➤ Exigem muito cuidado e tempo de preparação para garantir que todas as opções de respostas sejam oferecidas; ➤ Se alguma alternativa importante não foi previamente incluída, fortes vieses podem ocorrer, mesmo quando esteja sendo oferecida a alternativa "Outros. Quais?"; ➤ O respondente pode ser influenciado pelas alternativas apresentadas.
Dicotômicas	<ul style="list-style-type: none"> ➤ Rapidez e facilidade de aplicação, processo e análise; ➤ Facilidade e rapidez no ato de responder; ➤ Menor risco de parcialidade do entrevistador; ➤ Apresentam pouca possibilidade de erros; ➤ São altamente objetivas. 	<ul style="list-style-type: none"> ➤ Polarização de respostas e/ou possibilidade de forçar respostas em relação a um leque de opiniões; ➤ Podem levar a erros de medição, se o tema foi tratado de forma dicotômica, quando na verdade apresenta várias alternativas; ➤ Dependendo de como a pergunta é feita, questões com respostas dicotômicas são fortemente passíveis de erros sistemáticos.

2.2.1 Escalas

Quando se aplica um questionário fechado (múltipla escolha ou dicotômico) pretende-se medir aspectos como atitudes ou opiniões do público-alvo, e isso geralmente é possível com a utilização de escalas. As

escalas que se utilizam podem ser de quatro tipos: escala de Likert, VAS (Visual Analogue Scales), escala Numérica e escala Guttman.

A escala de Likert apresenta uma série de cinco proposições, das quais o respondente deve selecionar uma, podendo estas ser: concorda totalmente, concorda, sem opinião, discorda, discorda totalmente. É efetuada uma cotação das respostas que varia de modo consecutivo: +2, +1, 0, -1, -2 ou utilizando pontuações de 1 a 5. É necessária atenção quando a proposição é negativa. Nestes casos a pontuação atribuída deverá ser invertida.

VAS (Visual Analogue Scales) é um tipo de escala que advém da escala de Likert apresentando os mesmos objetivos, mas num formato diferente. Este tipo de escala baseia-se numa linha horizontal com 10 cm de comprimento apresentando nas extremidades duas proposições contrárias:

Útil _____ Inútil

O entrevistado deve responder à questão assinalando na linha a posição que corresponde à sua opinião.

A Escala Numérica deriva da escala anterior na qual a linha se apresenta dividida em intervalos regulares.

A escala de Guttman apresenta um conjunto de respostas que estão hierarquizadas. Deste modo se um respondente concordar com uma das opções esta concordância irá incluir todas as respostas que se encontram numa posição inferior na escala. Se o inquirido concordar com uma opção, mas não concordar com as anteriores, tal significará que a escala está mal construída. A cada item é atribuído cotação que se inicia em zero caso não seja escolhida nenhuma opção, um se for escolhida a primeira opção, dois se for escolhida a segunda opção e assim sucessivamente.

2.3 Decisões sobre a Formulação das Perguntas

Na formulação das perguntas deve-se cuidar para que as mesmas tenham o mesmo significado para o pesquisador e para o respondente, evitando-se assim um erro de medição. Sabe-se que a formulação tem efeito sobre as respostas. Esse efeito pode ser avaliado comparando-se os resultados em sub-amostras, de perguntas formuladas de forma diferente.

É conveniente fazer as seguintes recomendações sobre a formulação das perguntas:

- Usar comunicação simples e palavras conhecidas;
- Não utilizar palavras ambíguas.

Evitar:

- perguntas que sugiram a resposta;
- perguntas com conteúdo emocional e/ou sentimento de aprovação ou reprovação;
- referências a nomes que impliquem em aceitação ou rejeição ou tenham componente afetivo;
- alternativas implícitas;
- necessidade do respondente fazer cálculos para responder;
- perguntas de dupla resposta;
- alternativas longas;
- mudanças bruscas de temas, (fazer uma ligação entre os temas);
- contágio de respostas (efeito halo);
- vieses involuntários, motivados por reação visando prestígio por parte do respondente, retraimento defensivo diante de perguntas personalizadas e a atração exercida pela resposta positiva.

São condicionantes das respostas:

- Busca de conformidade ao grupo;
- Tendência de imitação social;
- Medo do julgamento do outro;
- Busca de prestígio social;
- Participação nas emoções coletivas;
- Submissão aos estereótipos culturais;
- Medo de mudanças.

Outros aspectos/questões referentes à redação das perguntas:

- A pergunta pode ser mal compreendida? Contém frases ou termos difíceis e/ou obscuros?
- Os termos utilizados serão bem compreendidos pelo público da pesquisa? Termos especializados são usados apenas quando realmente necessários, devendo-se assegurar que seu sentido se torne claro através de figuras ou de outros meios.
- A sentença é curta e simples? Sentenças longas e difíceis tendem a ser mal compreendidas.
- Existe indefinição ou ambigüidade? Qual o outro sentido que a pergunta poderá ter para quem responde?
- Enfatizar não intencionalmente uma palavra ou frase poderia mudar o sentido da pergunta?

Segundo Seltiz et al (1974) se, mesmo depois de certificar que as perguntas estão apresentadas da maneira mais clara possível, ainda existirem dúvidas quanto à compreensão, costuma-se incluir perguntas de acompanhamento, do tipo: "O que você quer dizer com isso?", "Você poderia exemplificar?". Dessa maneira, se torna possível verificar como o respondente entendeu a questão e o que pretendeu dizer.

- A pergunta exprime adequadamente todas as alternativas, ou mostra apenas um dos lados do tópico em questão? Ambos devem ser citados.
- O quadro de referência é claro e uniforme para todas as pessoas que respondem?
- A pergunta deixa claro que a pessoa deve respondê-la baseada naquilo que pensa ser a verdade e não naquilo que desejaria que fosse a verdade?
- A frase é enviesada, ou seja, está emocionalmente carregada ou deformada para determinado tipo de resposta?
- O que traria melhores resultados? Uma redação mais pessoal ou mais impessoal da pergunta?
- O que seria melhor? Apresentar a pergunta de maneira direta ou indireta? Não há recomendações concretas quanto ao emprego de perguntas indiretas. Suas possibilidades e limitações devem ser examinadas caso a caso, de acordo com o objetivo da pesquisa. Entram em questão problemas morais e técnicos: saber se as inferências pretendidas podem ser retiradas, sem prejuízo, das provas indiretas; se tal pergunta indireta irá enviesar a resposta; se as perguntas são altamente invasivas da intimidade etc. (Seltiz et al, 1974).

2.4 Decisões sobre a seqüência das perguntas

A ordem na qual as perguntas são apresentadas pode ser crucial para o sucesso da pesquisa. Não há regras estabelecidas, mas alguns cuidados devem ser tomados. Mattar (1994) recomenda:

- Iniciar o questionário com uma pergunta aberta e interessante (para deixar o respondente mais à vontade e assim ser mais espontâneo e sincero ao responder as perguntas restantes). Iniciar com perguntas sobre a opinião do respondente pode fazer com que se sinta prestigiado e se torne disposto a colaborar.

O primeiro contato do respondente com o questionário define sua vontade de respondê-lo ou até mesmo a decisão de não respondê-lo;

- Usar temas e perguntas gerais no início do questionário, deixando as perguntas específicas para depois (fechar o foco gradualmente);
- As perguntas mais pessoais, sensíveis ou embaraçosas devem ser feitas somente no final do questionário e convém que sejam alternadas com questões simples;
- Deve-se adotar uma ordem lógica de perguntas utilizando um fluxograma ou árvore de decisão para posicionar as perguntas;
- Dar uma seqüência lógica ao questionário. Mudanças de tópicos repentinas e "ir e voltar" ao assunto devem ser evitados;
- Informações que classificam social, econômica ou demograficamente o respondente devem ser colocadas no final, a não ser que alguma delas sirva como "filtro";
- Perguntas de caráter mais invasivo, ou que tratem temas delicados, não devem ser colocados no início do questionário e convém que sejam alternadas com questões simples;

Outra preocupação com o questionário é a de explicar as condições adequadas para o seu uso e aplicação, tanto no caso de formulários auto-preenchidos, quanto nos que utilizam entrevistadores. Devem ser fornecidas aos entrevistadores instruções claras de como proceder no campo, como abordar os

respondentes, como preencher os instrumentos, etc. A seguir, são apresentados alguns pontos sobre os quais os entrevistadores devem ser orientados.

- Proporcionar ao respondente uma situação de liberdade, em que a pessoa seja estimulada a apresentar francamente suas opiniões;
- Garantir, se for o caso, o anonimato do respondente;
- O entrevistador deve ser educado, amigável e imparcial;
- Nunca deverá mostrar surpresa ou desaprovação diante das opiniões de quem responde;
- As perguntas precisam ser apresentadas da maneira exata, com as mesmas palavras que foram propostas;
- Qualquer explicação improvisada da pergunta é proibida. Em casos em que se imagine, de antemão, que surgirão dúvidas, esclarecimentos devem ser previamente elaborados;
- As perguntas devem seguir a ordem exata em que aparecem no questionário;
- O entrevistador deve apresentar todas as perguntas, e jamais responder alguma por dedução própria;
- Espera-se que o entrevistador registre fiel e integralmente a resposta;
- É necessário que os entrevistadores sejam orientados em relação ao processo de amostragem. Por exemplo, como proceder em casos de recusas ou ausências.

2.5 Decisões sobre a apresentação e o lay-out do questionário (características físicas)

São pontos a serem definidos nesta fase: número de páginas; qualidade do papel e da impressão; tipos e tamanho de letras; posicionamento e tamanho dos espaços entre questões; cores da tinta e do papel para as respostas; espaço para resposta de cada questão; separação de campos para facilidade de digitação (praticamente obrigatória para se compilar as respostas e processá-las em tempo reduzido); impressão em frente e verso ou só na frente.

Tais itens são relevantes para se ganhar a colaboração dos respondentes. Quanto melhor e mais adequada for a apresentação, maior a probabilidade de se elevar o índice de respostas.

2.6 Decisões quanto ao Pré-teste

É importante a realização de um pré-teste porque é provável que não se consiga prever todos os problemas e/ou dúvidas que podem surgir durante a aplicação do questionário. Sem o pré-teste, pode haver grande perda de tempo, dinheiro e credibilidade caso se constate algum problema grave com o questionário já na fase de aplicação. Nesse caso o questionário terá que ser refeito e estarão perdidas todas as informações já colhidas.

Segundo Mattar (1994), os pré-testes podem ser realizados inclusive nos primeiros estágios, quando o instrumento ainda está em desenvolvimento, quando o próprio pesquisador pode realizá-lo, através de entrevista pessoal. O pré-teste é, segundo Goode e Hatt (1972), um ensaio geral. Cada parte do procedimento deve ser projetada e implementada exatamente como o será na hora efetiva da coleta de dados. As instruções para a entrevista devem estar na formulação final, e serem obedecidas rigorosamente, para se ver se são ou não adequadas. O questionário deve ser apresentado na forma final e a amostra (embora menor) deve ser obtida segundo o mesmo plano que gerará a amostra final. Os resultados do pré-teste são então tabulados para que se conheçam as limitações do instrumento. Isto incluirá a proporção de respostas do tipo "não sei", de questões difíceis, ambíguas e mal formuladas, a proporção de pessoas que recusam a entrevista, bem como os comentários feitos pelos respondentes sobre determinadas questões.

Caso o pré-teste revele necessidade de muitas alterações, o questionário revisado deverá ser então novamente testado. O processo será repetido tantas vezes quantas forem necessárias, até que o instrumento se encontre maduro, pronto para ser aplicado. De acordo com Mattar (1994), para instrumentos que foram cuidadosamente desenvolvidos, dois ou três pré-testes costumam ser suficientes.

3. Populações e amostras

Uma noção importante referente à amostragem é reconhecer que uma amostra válida simplesmente não é o conjunto de respostas que adquirimos quando administrarmos um questionário. Um conjunto de respostas só uma amostra válida, em condições estatísticas, se for obtido por um processo de amostragem aleatório. (Kitchenham e Pflieger, 2002a).

3.1 Amostras e populações

Para obter uma amostra, você tem que começar definindo uma população alvo. A população alvo é o grupo ou os indivíduos a quem a pesquisa se aplica. Idealmente, uma população alvo deveria ser representada por uma lista finita de todos seus membros. Por exemplo, quando entrevistadores pesquisam membros do público sobre suas preferências de votação, eles usam a lista eleitoral como sua lista de população alvo.

Uma amostra válida é um subconjunto representativo da população alvo. A palavra crítica em nossa definição de amostra é a palavra "representante." Se não tivermos uma amostra representativa, não podemos declarar que os resultados generalizam à população alvo. Se nossos resultados não generalizarem, eles têm pouco valor. Assim, a principal preocupação quando amostramos uma população é assegurar que essa amostra é representativa.

3.2 Obtendo uma amostra válida

Não se pode amostrar uma população se não pudermos especificar o que aquela população é. Nossa avaliação inicial da população alvo deveria surgir dos objetivos da pesquisa, não de um senso de quem está disponível para responder nossas perguntas. Quanto mais claros os objetivos, mais fácil será definir a população alvo. A população alvo específica pode ser um subconjunto de uma população maior.

É freqüentemente instrutivo considerar a população alvo e procedimento de amostragem do ponto de vista de análise de dados. Podemos fazer isso durante o projeto do questionário, mas também deveríamos reavaliar a situação depois de qualquer pré-teste do instrumento de pesquisa. Nesse momento teremos algumas respostas atuais, assim podemos experimentar nossos procedimentos de análise. Precisamos considerar se as análises conduzirão a qualquer conclusão significante, em particular:

- Os resultados de análise atendem aos objetivos de estudo?
- A população alvo pode responder nossas perguntas de pesquisa?

3.2.1 Métodos de Amostragem

Uma vez definida a população alvo apropriada, devemos usar um método rigoroso de amostragem. Se quisermos fazer inferências fortes à população alvo, necessitamos de um método de amostragem probabilístico.

São descritos a seguir uma variedade de métodos de amostragem, probabilísticos e não-probabilísticos.

3.2.2 Métodos de amostragem probabilísticos

Uma amostra é probabilística quando qualquer membro de uma população alvo tem uma probabilidade conhecida (> 0) de ser incluído na amostra. O objetivo de uma amostra probabilística é eliminar a subjetividade e obter uma amostra que seja imparcial e representativa da população alvo. É importante lembrar que não podemos fazer nenhuma conclusão estatística dos dados obtidos a menos que tenhamos uma amostra probabilística.

3.2.3 Amostra aleatória simples

Uma amostra aleatória simples é quando qualquer membro da população alvo tem a mesma probabilidade de ser incluído na amostra. Há uma variedade de modos de selecionar uma amostra aleatória de uma lista da população. Um modo é usar um gerador de número aleatório para atribuir um número aleatório a cada membro da população alvo, ordenar os membros da lista de acordo com o número aleatório e escolher os primeiros n membros da lista onde n é o tamanho de amostra.

3.2.4 Amostra aleatória estratificada

Neste caso, a população designada é dividida em subgrupos chamada estratos. Cada estrato é amostrado separadamente. Estratos são usados quando nós esperamos segmentos diferentes da população alvo para responder diferentemente a nossas perguntas, ou quando esperamos que diferentes segmentos da população tivessem tamanhos diferentes. Por exemplo, nós podemos estratificar uma população

designada em base de sexo, porque homens e mulheres geralmente respondem de forma diferente a questionários.

O número de membros selecionados de cada estrato geralmente é proporcional ao tamanho do estrato. Em uma pesquisa de engenharia de software, temos freqüentemente bem menos mulheres que homens em nossa população alvo, assim podemos amostrar usando estratos para assegurar que temos um número apropriado de respostas de mulheres.

Amostras aleatórias estratificadas são úteis para populações não-homogêneas, mas elas são complicadas para analisar do que a amostras aleatórias simples.

3.2.5 Amostragem sistemática

Amostragem sistemática envolve selecionar todo enésimo membro da lista de uma população. Se a lista é aleatória, enquanto selecionar todo enésimo membro é outro método de se obter uma amostra aleatória simples. Porém, se a lista não for aleatória, este procedimento pode introduzir parcialidade. Ordem não-aleatória incluiria ordem alfabética ou data de nascimento.

3.2.6 Amostragem baseada em agrupamento

Amostragem baseada em agrupamento é o termo utilizado para questionar indivíduos que pertencem a grupos definidos. Por exemplo, podemos questionar todos os membros de um grupo familiar ou todos os pacientes em hospitais específicos. Procedimentos de randomização estão baseados no agrupamento, não no indivíduo. Esperamos que os membros de cada agrupamento dêem respostas mais semelhantes do que membros de agrupamentos diferentes. Quer dizer, são esperadas respostas correlatas dentro de um agrupamento. Há métodos bem definidos para analisar dados de agrupamento, mas a análise é mais complexa do que uma amostra aleatória simples.

3.3 Métodos de amostragem não-probabilísticos

São criadas amostras não-probabilísticas quando os respondentes forem escolhidos porque eles são facilmente acessíveis ou os pesquisadores têm alguma justificativa por acreditar que eles são representativos da população. Este tipo de amostra corre o risco de ser parcial, assim é perigosa para se tirar qualquer conclusão importante.

Existem três razões para se usar amostras não-probabilísticas:

1. É difícil identificar a população alvo. Por exemplo, se quisermos inspecionar os hackers de software, eles podem ser difíceis de serem encontrados.
2. A população designada é muito específica e de disponibilidade limitada. Por exemplo, se quisermos inspecionar os executivos de companhias que empregam mais de 5000 engenheiros de software, pode não ser possível confiar em uma amostra aleatória. Podemos ser forçados a questionar apenas os executivos que estão dispostos a participar.
3. A amostra é um estudo piloto, não a pesquisa final, e um grupo não aleatório está disponível.

3.3.1 Amostragem de conveniência

Amostragem de conveniência envolve obter respostas de pessoas que estão disponíveis e dispostas a participar. O problema principal desta abordagem é que as pessoas que estão dispostas a participar podem diferir de modo importante dos que não estão dispostos. Vemos esse tipo de amostragem particularmente em sites da Web onde as pessoas que têm reclamações são mais prováveis de prover avaliação do que as que estão satisfeitas com um produto ou serviço.

3.3.2 Amostragem bola de neve (SnowBall)

Envolve pedir às pessoas que participam de uma pesquisa para nomear outras pessoas que estariam dispostas a participar. A amostragem continua até que o número exigido de respostas seja obtido. Esta técnica é freqüentemente usada quando a população for de difícil identificação pelos pesquisadores. Por exemplo, poderíamos esperar que os hackers de software se conheçam entre si, assim se achássemos um hacker para participar da pesquisa, poderíamos lhe pedir que identificasse outros possíveis participantes.

3.3.3 Amostragem de cota

Amostragem de cota é a versão não-probabilística da amostragem aleatória estratificada. A população alvo é dividida em estratos apropriados baseados em subgrupos conhecidos (por exemplo. sexo, formação educacional, tamanho de companhia etc.). Cada estrato é amostrado (usando conveniência ou técnicas de bola de neve) de forma que número de respondentes em cada subgrupo corresponde a sua proporção na população.

3.3.4 Grupos de foco

Grupos de foco normalmente são formados pelos pesquisadores a partir de seus contatos pessoais. Eles normalmente consistem em 10 a 20 pessoas que representam alguma população. Grupos de foco são geralmente usados em estudos piloto de pré-pesquisa.

3.4 Tamanho de amostra

Uma grande preocupação é determinar um tamanho de amostra apropriado. Há duas razões por que tamanho da amostra é importante. Primeira, um tamanho de amostra inadequado pode conduzir a resultados que não são estatisticamente significantes. Em outras palavras, se o tamanho de amostra não for bastante grande, não iremos obter uma conclusão razoável e não podemos generalizar à população alvo. Um exemplo extremo deste problema é obtermos uma única resposta; não podemos tirar nenhuma conclusão de um único respondente. Segunda, amostragem inadequada de agrupamentos ou estratos incapacita nossa habilidade de comparar e contrastar subconjuntos diferentes da população.

Para determinar um tamanho de amostra adequado ou mínimo, precisamos saber quatro coisas sobre nosso estudo:

1. O nível alfa que pretendemos usar, onde alfa é a probabilidade de um erro Tipo I (quer dizer, a probabilidade de falsamente rejeitar a hipótese inválida). Alfa é normalmente fixa em 0.05 ou 0.01.
2. O nível de beta que nós pretendemos usar, onde beta é a probabilidade de um erro Tipo II (quer dizer, a probabilidade de falsamente aceitar a hipótese inválida). Beta é normalmente fixa em 0.20. Falamos freqüentemente sobre o poder de um teste ou experiência; poder é calculado como 1-beta. O poder de um teste é a probabilidade de aceitar corretamente a hipótese alternativa.
3. O tamanho de efeito que é a diferença de resultados entre dois grupos. Por exemplo, suponha que queremos investigar se há diferenças de pagamento entre engenheiros de software de diferentes sexos. Poderíamos pesquisar os homens e mulheres que se formaram em 1998 e perguntar qual é o salário base deles. O tamanho de efeito é a diferença entre o salário masculino comum e o salário feminino comum.
4. A variância de efeito é o grau de variação dos dados dentro de um grupo. Usando o exemplo do salário, podemos obter a variância de valores de salário para os homens e para as mulheres.

Claro que, o tamanho de efeito e de variância é o que nós esperamos obter como resultado de nossa pesquisa, assim precisamos de informação anterior para determinar um tamanho de amostra apropriado. Podemos obter tal informação de pesquisas prévias, pesquisa piloto ou opinião de um especialista.

No caso simples de avaliar o tamanho de amostra, assumindo uma distribuição Normal para as variáveis de resposta, dois grupos com números iguais em cada grupo e variâncias iguais dentro do grupo, o tamanho de amostra (por grupo) é:

$$\left[\frac{(z_a - z_b) * \sigma}{\mu_1 - \mu_2} \right]^2$$

onde

$\mu_1 - \mu_2$ é o tamanho de efeito.

σ é a variância padrão.

z_a é a parte superior na distribuição normal padrão que corresponde a α . Por exemplo, $z_a = 1.96$ se $\alpha = 0.05$.

z_b é a parte inferior na distribuição normal padrão que correspondente a β . $z_b = -0.84$ se $\beta = 0.20$.

Não devemos esquecer os não respondentes quando fixarmos o tamanho da amostra. Por exemplo, se a fórmula nos fala que o tamanho de amostra ótima é 50, mas esperamos uma taxa de resposta de 80%, devemos aumentar o tamanho da amostra para 63.

4. Análise de Dados

Neste ponto, assumimos que você já projetou e administrou a sua pesquisa e agora você está pronto para analisar os dados coletados. Agora abordaremos alguns assuntos gerais envolvendo análise de dados de pesquisa. Porém, não iremos descrever como analisar todos os tipos de dados de pesquisa em detalhes, desse modo iremos nos concentrar na discussão de alguns erros de análise mais comuns e como evitá-los. (Kitchenham e Pfleeger, 2002c)

4.1 Validação de dados

Antes de empreender alguma análise detalhada, deveriam ser verificadas a consistência e a integridade das respostas. É importante ter uma política para controlar questionários incoerentes ou incompletos. Se a maioria dos respondentes atendeu a todas as perguntas, podemos tomar a decisão de rejeitar questionários incompletos. Porém, temos que investigar as características dos questionários rejeitados da mesma forma que nós investigamos não-respostas para assegurar que não introduzimos nenhuma parcialidade sistemática. Alternativamente, podemos achar que a maioria dos respondentes se omitiu em algumas perguntas específicas. Neste caso, é mais apropriado remover essas perguntas da análise.

Às vezes podemos usar todos os questionários, até mesmo quando alguns estão incompletos. Neste caso, temos tamanhos de amostra diferentes para cada pergunta que analisamos e temos que nos lembrar de informar o tamanho de amostra atual para cada amostra estatística. Esta aproximação é satisfatória para análises como cálculo de amostras estatísticas ou comparação de valores médios, mas não para correlação ou estudos de regressão. Sempre que a análise envolve duas ou mais perguntas ao mesmo tempo, precisamos definir um procedimento para tratar valores perdidos.

Em alguns casos, é possível usar técnicas estatísticas para "imputar" os valores de dados perdidos (Little e Rubin, 1987). Porém, tais técnicas são normalmente impróprias quando a quantidade de dados perdidos for excessiva e/ou os valores são categóricos em vez de numéricos.

É importante para reduzir a chance de questionários incompletos quando nós projetarmos e testamos nossos instrumentos. Uma justificativa muito forte para pesquisas de piloto é poder detectar aquelas perguntas enganosas e/ou instruções incompletas antes de realizar a pesquisa principal.

Precisamos freqüentemente dividir nossas respostas em subgrupos mais homogêneos antes da análise. A divisão é normalmente baseada em informação demográfica. Podemos comparar as respostas obtidas de subgrupos diferentes ou simplesmente informar os resultados separadamente para cada subgrupo. Em alguns casos, a divisão pode ser usada para aliviar alguns erros do projeto inicial.

4.2 Codificação de dados

Às vezes é necessário converter dados de escalas nominais e ordinais de nomes de categoria para contagens numéricas antes de serem inseridos nos arquivos de dados. A idéia da tradução não é analisar dados de escala nominais e ordinais como se eles fossem valores numéricos simples. Isso é realizado porque muitos pacotes estatísticos não suportam categorias representadas por uma seqüência de caracteres. Em muitos casos, são colocados códigos no questionário junto com nomes de categoria, desse modo a codificação é realizada durante o projeto ao invés de ser realizada durante análise de dados.

Um problema de codificação mais complexo surge para perguntas abertas. Neste caso, categorias de resposta precisam ser construídas depois que os questionários forem devolvidos. Identificar se duas respostas diferentes são equivalentes ou não exige perícia humana. Nesses casos, é sábio pedir a várias pessoas diferentes que codifiquem respostas e comparem os resultados, de forma que a parcialidade não seja introduzida pela categorização.

4.3 Análise de dados padrão

A análise de dados específica, que você precisa, depende do projeto de pesquisa e o tipo de escala de respostas (nominal, ordinal, intervalar, etc.). As estatísticas de população mais comuns para valores numéricos são:

$$\text{População Total: } X_T = \sum_{i=1}^N X_i$$

$$\text{Média da População: } \hat{X} = \left(\sum_{i=1}^N X_i \right) / N$$

$$\text{Variância da População: } \sigma_x^2 = \sum (X_i - \hat{X}) / N$$

onde N é o tamanho da população.

Para variáveis de dicotômicas (Sim/Não ou 0/1), as estatísticas de população mais comuns são:

$$\text{Proporção: } P_Y = \left(\sum_{i=1}^N Y_i \right) / N$$

$$\text{Variância de uma proporção: } \sigma_Y^2 = P_Y(1 - P_Y)$$

onde Y_i é uma variável dicotômica que leva o valor 1 ou 0 e N é o tamanho da população.

Se tivermos uma amostra aleatória de tamanho n obtida de uma população de tamanho N, podemos calcular as estatísticas de população de nossa amostra utilizando as seguintes equações:

$$\text{Total: } \hat{X}_T = \left(\frac{N \sum_{i=1}^n x_i}{n} \right)$$

$$\text{Média: } \bar{x} = \left(\sum_{i=1}^n x_i \right) / n$$

$$\text{Variância: } \sigma_x^2 = \left(\frac{N-1}{N} \right) (s_x^2) \quad \text{onde: } s_x^2 = \left(\sum_{i=1}^n (x_i - \bar{x})^2 \right) / (n-1)$$

$$\text{O erro padrão da estimativa do total é: } N \left(\frac{s_x}{n} \right) \sqrt{\frac{N-n}{N}}$$

$$\text{O erro padrão da estimativa da média é: } \left(\frac{s_x}{n} \right) \sqrt{\frac{N-n}{N}}$$

$$\text{Para a proporção } \hat{P}_Y = \frac{\sum_{i=1}^n y_i}{n} \text{ e o erro padrão da estimativa da proporção é: } \sqrt{\frac{N-n}{N}} \sqrt{\frac{P_Y(1-P_Y)}{n-1}}$$

Estas podem não ser as fórmulas que você estaria esperando ao ler um livro básico em estatística e análise de dados. Os erros padrões incluem o termo $\sqrt{(N-n)/N}$ que será chamado de correção de população finita (fpc)(Levy e Lemeshow, 1999). O fpc pode ser re-escrito como $\sqrt{1-(n/N)}$ onde se pode observar que se N tender ao infinito, o fpc se aproxima de 1 e as fórmulas de erros padrão se aproxima das fórmulas habituais. Se $n = N$, as condições de erro padrão são zero porque os valores da média, total e de proporção são conhecidos e portanto não sujeito a erro.

Nota. A equação que nós demos por determinar tamanho de amostra ignora o fpc, e se usada como está resultará em um resultado sobreestimado do tamanho de amostra exigida. Porém, é melhor para ter mais do que menos em uma amostra.

4.4 Analisando dados ordinais e nominais

A análise de dados numéricos é relativamente direta. Porém, há problemas adicionais se seus dados forem ordinais ou nominais.

4.4.1 Dados ordinais

Um grande número de pesquisas solicita que as pessoas respondam a perguntas numa escala ordinal. Por exemplo, são pedidos aos respondentes que especifiquem até que ponto eles concordam com uma determinada declaração. Aos respondentes são oferecidas as seguintes opções: concordo totalmente, concordo, não concordo nem discordo, discordo ou discordo totalmente. É uma prática comum converter a escala ordinal em um valor numérico correspondente (por exemplo: números 1 a 5) e analisar os dados como se eles fossem dados numéricos simples. Há ocasiões em que essa abordagem é razoável, mas viola as regras matemáticas para a análise de dados ordinais. Usando uma conversão de ordinal para vínculos numéricos um risco que análise subsequente dará resultados enganosos.

Figura 1 representa uma pesquisa com três perguntas, cada uma com uma escala ordinal de 5 opções de resposta (rotuladas SP1, SP2, SP3, SP4, SP5) onde temos 100 respondentes. A Figura 3 apresenta as respostas relacionadas a cada ponto da escala; por exemplo, para pergunta 1, os respondentes escolheram SP1 10 vezes, SP2 20 vezes, SP3 40 vezes, SP4 20 vezes e SP5 10 vezes. Se convertemos o que a escala aponta para um equivalente numérico (de 1 a 5) e determinar o valor médio, chegaremos a uma média igual a 3 para todas as três perguntas. Porém, não podemos concluir que todas as respostas são equivalentes. No caso da pergunta 1, temos uma distribuição simétrica com um único pico. Isso pode ser considerado como próximo de uma distribuição Normal com um valor médio de 3. No caso de pergunta 2, temos uma distribuição bimodal. Para distribuições de bimodal, os dados não são Normais. Além disso, não há nenhuma tendência central, assim a média não é uma estatística útil. No caso da terceira pergunta, temos um número igual de respostas em cada categoria, típico de uma distribuição uniforme. Uma distribuição uniforme não tem nenhuma tendência central, e novamente o conceito de média não é útil.

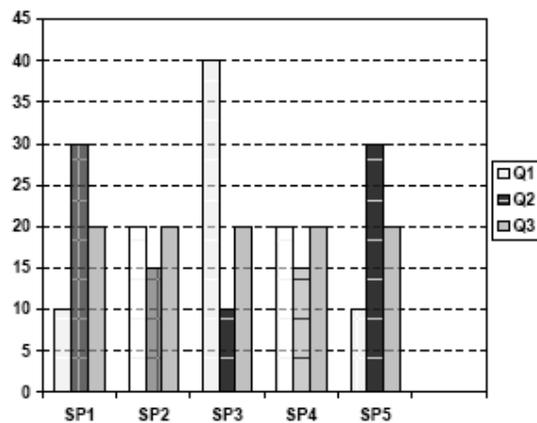


Figura 3 - Respostas a três perguntas com uma escala ordinal de 5 pontos

Em geral, se os dados possuírem um único pico e se aproximarem de uma Normal, os riscos de uma análise incorreta são baixos se convertermos em valores numéricos. Porém, devemos considerar se tal conversão é necessária. Há três abordagens que podem ser usadas se quisermos evitar violações de escala:

1. Podemos usar as propriedades da distribuição multinomial para calcular a proporção da população em cada categoria e então determinar o erro padrão da estimativa.
2. Podemos talvez converter uma escala ordinal em uma variável dicotômica. Por exemplo, se estamos interessados em comparar se a proporção de quem concorda ou concorda totalmente é maior em um grupo que outro, podemos re-codificar nossas respostas em uma variável dicotômica (por exemplo, podemos codificar "concordo totalmente" ou "concordo" como 1 e todas as outras respostas como 0) e usar as propriedades da distribuição binomial. Essa técnica também é útil se quisermos avaliar o impacto de outras variáveis em uma variável de escala ordinal.
3. Podemos usar a correlação de postos de Spearman ou o tau de Kendall (Siegel e Castellan, 1998) para medir a associação entre variáveis de escala ordinal.

Há duas ocasiões onde não há nenhuma alternativa real a violações de escala:

1. Se quisermos avaliar a confiança de nosso instrumento de pesquisa usando a estatística de alfa de Cronbach (Cronbach, 1951).
2. Se quisermos adicionar conjuntamente medidas de escala ordinais de variáveis relacionadas, para gerar contagens totais para um conceito.

Porém, em ambos os casos, se não tivermos uma resposta próxima de uma Normal, os resultados da análise dos dados podem estar errados. É importante entender o tipo de escala utilizado em nossos dados e analisá-los adequadamente.

4.4.2 Dados nominais

A forma mais comum de análise aplicada a dados nominais é determinar a proporção de respostas em cada categoria. Assim, a menos que haja só duas categorias, não há nenhuma escolha exceto o uso das propriedades da distribuição multinomial (ou possivelmente o hipergeométrico) para determinar erros padrões das proporções. Porém, ainda é possível usar tabelas multi-modo (multi-way) mesas e testes chi-quadrados para medir associações entre variáveis de escalas nominais (veja (Siegel e Castellan, 1998), Seção 9.1).

5. Conclusões

Espera-se que ao longo deste trabalho o leitor tenha compreendido a importância do planejamento na realização de um questionário. Todas as indagações quanto ao conteúdo, forma, redação e seqüência devem ser feitas para cada questão. Esse planejamento não se limita ao desenvolvimento das questões, mas também a realização de pré-testes, determinação da população alvo e da amostra, e análise dos dados. Tudo isso deve ser realizado antes de se questionar toda a amostra com o propósito de não perder todo trabalho empenhado.

6. Referências

AAKER, ET AL (2001) "Marketing Research" (7th Ed.), New York: John Wiley & Sons, Inc

CRONBACH, L. J. (1951) Coefficient alpha and internal structure of tests, *Psychometrika*, 16(2), pp. 297-334.

GOODE, Willian J.; HATT, Paul K. (1972) *Métodos em Pesquisa Social*. 4a ed. São Paulo: Nacional.

KITCHENHAM, B. and PFLEEGER, S. (2002a). Principles of survey research: part 2: designing a survey. *ACM SIGSOFT Software Engineering Notes*, 27(1):44–45.

KITCHENHAM, B. and PFLEEGER, S. (2002c). Principles of survey research: part 6: data analysis. *ACM SIGSOFT Software Engineering Notes*, 28(2):24–27.

LEVY, P. S. and LEMESHOW, S. (1999) Sampling and Populations, Third Edition, Wiley Series in Probability and Statistics, John Wiley & Sons, New York.

LITTLE, R. J. A. and RUBIN, D. B. (1987) Statistical Analysis with Missing Data, Wiley, New York.

MATTAR, F. N. (1994) Pesquisa de marketing: metodologia, planejamento, execução e análise, 2a. ed. São Paulo: Atlas, 2v., v.2.

SELLTIZ, Claire et al. (1974) Métodos de pesquisa nas relações sociais. 3 a . ed. São Paulo: E.P.U.

SIEGEL, S. and CASTELLAN, N. J (1998) Nonparametric Statistics for the Behavioral Sciences, 2nd Edition, McGraw-Hill Book Company, N.Y.