

EBSERH

HOSPITAIS UNIVERSITÁRIOS FEDERAIS

PESQUISA DE SATISFAÇÃO DOS USUÁRIOS DE HOSPITAIS UNIVERSITÁRIOS FEDERAIS

(1º CICLO 2016)

BRASÍLIA, MARÇO DE 2016

Presidente

Newton Lima Neto

Diretora Vice-Presidente Executiva:

Jeanne Liliane Marlene Michel

Diretor de Atenção à Saúde

Adriana Karla Nunes Barbuio

Diretor de Administração e Infraestrutura

Garibaldi Jose Cordeiro de Albuquerque

Diretor de Gestão de Pessoas

Diretor substituto: Marcos Aurélio Souza Brito

Diretoria de Controladoria e Finanças

Cristian de Oliveira Lima

Diretoria de Gestão de Processos e Tecnologia da Informação

Cristiano Cabral

COORDENAÇÃO GERAL**Ouvidor Geral Ebserh**

Josué Fermon Ribeiro

COORDENAÇÃO TÉCNICA**Ouvidoria Hospital Universitário de Brasília/Representante Ouvidoria Geral**

Paulo Rodrigo Ribeiro Guimarães

Ouvidoria Hospital Universitário da Universidade Federal do Maranhão

Pedro Germano Nobre Neto

Ouvidoria Hospital Universitário da Universidade Federal do Piauí

Leila Leal Leite

Ouvidoria Hospital Universitário de Dourados da Universidade Federal de Dourados

Edméia Barrios de Azambuja Gonçalves

Ouvidoria Hospital Universitário Júlio Muller da Universidade Federal de Mato Grosso

Jovanildes de Fátima Silva

Ouvidoria Hospital Universitário Onofre Lopes da Universidade Federal do Rio Grande do Norte

Nelma Gomes Bezerra

Hospitais participantes:

1. Complexo Hospital de Clínicas da Universidade Federal do Paraná
2. Hospital das Clínicas da Universidade Federal de Goiás
3. Hospital das Clínicas da Universidade Federal de Minas Gerais
4. Hospital de Clínicas da Universidade Federal de Pernambuco
5. Hospital Escola da Universidade Federal do Triângulo Mineiro
6. Hospital Regional João Batista de Carvalho Daltro da Universidade Federal de Sergipe
7. Hospital Universitário Alberto Antunes da Universidade Federal de Alagoas
8. Hospital Universitário Ana Bezerra da Universidade Federal do Rio Grande do Norte
9. Hospital Universitário Cassiano Antônio de Moraes da Universidade Federal do Espírito Santo
10. Hospital Universitário da Universidade Federal de Grande Dourados
11. Hospital Universitário da Universidade Federal de Juiz de Fora
12. Hospital Universitário da Universidade Federal de Pelotas
13. Hospital Universitário da Universidade Federal de Santa Maria
14. Hospital Universitário da Universidade Federal de São Carlos
15. Hospital Universitário da Universidade Federal de Sergipe
16. Hospital Universitário da Universidade Federal do Maranhão
17. Hospital Universitário da Universidade Federal do Piauí
18. Hospital Universitário da Universidade Federal do Vale São Francisco
19. Hospital Universitário de Brasília da Universidade de Brasília
20. Hospital Universitário Doutor Miguel Riet Corrêa Junior da Universidade Federal do Rio grande
21. Hospital Universitário Getúlio Vargas da Universidade Federal do Amazonas
22. Hospital Universitário Júlio Bandeira da Universidade Federal de Campina Grande
23. Hospital Universitário Júlio Muller da Universidade Federal de Mato Grosso
24. Hospital Universitário Lauro Wanderley da Universidade Federal da Paraíba
25. Hospital Universitário Maria Aparecida Pedrossian da Universidade Federal de Mato Grosso do Sul
26. Hospital Universitário Onofre Lopes da Universidade Federal do Rio Grande do Norte
27. Hospital Universitário Professor Edgard Santos da Universidade Federal da Bahia
28. Hospital Universitário Walter Cantídio da Universidade Federal do Ceará
29. Maternidade Escola Assis Chateaubriand da Universidade Federal do Ceará
30. Maternidade Escola Climério de Oliveira da Universidade Federal da Bahia
31. Maternidade Escola Januário de Cicco da Universidade Federal do Rio Grande do Norte

I. Objetivo

A Pesquisa de Satisfação dos Usuários dos Hospitais Universitários Federais é uma iniciativa da Ouvidoria Geral da Empresa Brasileira de Serviços Hospitalares Ebserh e tem por objetivo avaliar a percepção de satisfação dos usuários dos hospitais sob gestão da Ebserh, com vistas ao aprimoramento da qualidade do atendimento.

II. Justificativa

Segundo o Ministério da Saúde (BRASIL, 2005), satisfação do usuário é a “condição que se percebe nos pacientes, no âmbito da saúde, consubstanciada em um contentamento advindo da realização de suas expectativas em relação à atenção sanitária recebida.” De um modo geral, embora não haja um referencial teórico estável e consensual sobre esse conceito (ESPERIDIÃO & TRAD, 2006), a satisfação dos usuários é considerada um aspecto central para a avaliação de serviços e sistemas de saúde.

Nos Hospitais Universitários Federais, três requisitos sustentam a necessidade de realização desse tipo de pesquisa:

- No âmbito da administração pública federal, a aplicação desse instrumento está prevista no Decreto nº 6.932/2009, para assegurar a participação do cidadão na avaliação dos serviços prestados.
- De parte das recomendações dos órgãos de controle, dentre as medidas relacionadas ao processo de reestruturação dos Hospitais Universitários Federais no Acórdão nº 2.813/2009 do Tribunal de Contas da União, consta a previsão de mecanismos de avaliação periódica, inclusive pesquisas de satisfação, e de divulgação de resultados.
- Sob o ponto de vista institucional, espera-se que a realização da pesquisa na etapa de assunção, pela Ebserh, dos hospitais contratantes se constitua como referencial para a análise dessa categoria e também para subsidiar a formulação e execução de ações nesse campo.

III. Estratégias de execução

III.1. Universo da Pesquisa

A Pesquisa será aplicada em todos os hospitais universitários cuja gestão é de responsabilidade da Empresa Brasileira de Serviços Hospitalares (Ebserh). A lista completa desses hospitais encontra-se na página 02 deste documento.

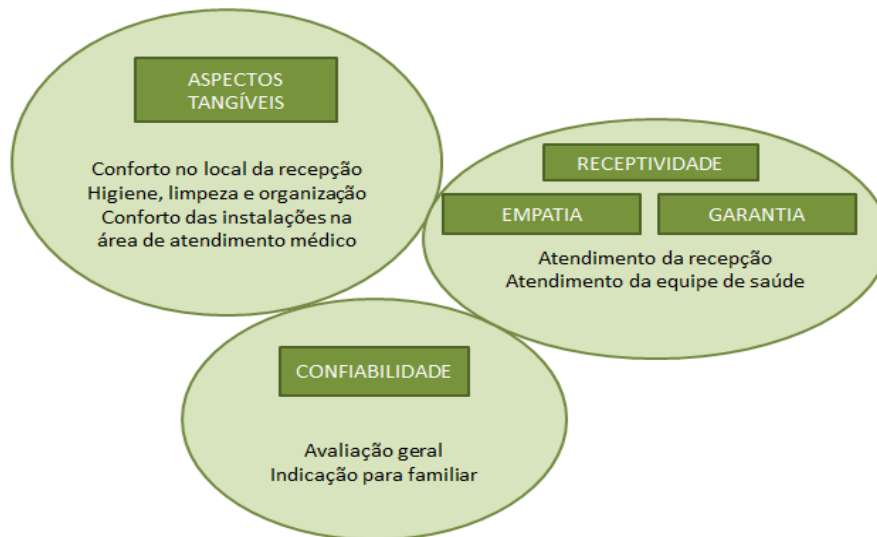
III.2. Instrumento de pesquisa: questionário estruturado

A opção por um questionário estruturado (Anexo I) se deveu às vantagens que apresenta tanto sob o ponto vista de aplicação, por sua simplicidade e baixo custo, quanto na análise, pela possibilidade de padronizar as informações, o que facilita a análise. Alguns critérios foram adotados na elaboração das questões e categorias de resposta:

- O questionário deve ser sintético. Estabeleceu-se que teria, no máximo, dez perguntas.
- As perguntas são claras para os entrevistados e relevantes para a gestão.
- Cada questão é aplicável a todos os grupos de usuários, de forma que todas as perguntas sejam respondidas por qualquer entrevistado.
- As perguntas têm conteúdo semânticos diferentes, de forma a evitar ambiguidades e duplicidades.

O questionário avalia aspectos relacionados à estrutura, atendimento e avaliação geral do hospital, que apontam questões relacionadas a cinco dimensões da prestação de serviços passíveis de avaliação por parte dos públicos alvos (BRASIL, 2013), de acordo com a Figura 1.

Figura 1. Aspectos utilizados e dimensões relacionadas no questionário da Pesquisa de Satisfação dos Usuários de Hospitais Universitários Federais (Empresa Brasileira de Serviços Hospitalares julho de 2013).



Aspectos tangíveis: Dizem respeito à avaliação das instalações e materiais usados durante a provisão do serviço: aparência física das instalações, limpeza e conservação dos equipamentos, aparência do pessoal, materiais usados, conforto dos móveis, das salas de espera, dos escritórios, etc.

Confiabilidade: refere-se à capacidade da instituição executar seus serviços de forma confiável e precisa, de acordo com o que foi estabelecido ou prometido.

Receptividade: corresponde ao interesse e vontade do atendente em auxiliar e atender os públicos alvos.

Garantia: conhecimento e cortesia dos funcionários de uma organização e sua capacidade de inspirar confiança nos públicos alvos. Essa dimensão agregou quatro das características iniciais da prestação de serviços: competência, cortesia, credibilidade e segurança.

Empatia: corresponde ao cuidado e atenção individualizados dados pelos funcionários aos públicos alvos. A empatia inclui três das características iniciais da prestação de serviços: acesso, comunicação e entendimento.

III.3. Aplicação do questionário

Com relação à realização das entrevistas, o responsável para a execução da Pesquisa é a Ouvidoria de cada HU, HC ou Maternidade, por seu papel de representação e mediação dos cidadãos e pelo caráter pedagógico, instrumental e estratégico de sua atuação. Serão entrevistados usuários nas áreas de Internação e Ambulatorial.

Com relação ao tipo de amostra, será utilizada a amostragem por proporções, utilizada para a determinação do tamanho da amostra para populações finitas.

Com relação à periodicidade, a pesquisa será aplicada em dois ciclos no ano vigente, conforme cronograma estabelecido pelos Ouvidores por meio de videoconferências organizadas pelo Grupo de Coordenação da Pesquisa de Satisfação dos Usuários dos HU Federais ligados à EBSERH, representante da Ouvidoria Geral neste processo.

III.4. Questões sobre a abordagem dos usuários

Destacam-se alguns aspectos sobre a estratégia de comunicação com o usuário durante a coleta de dados (BRASIL, 2013), uma etapa fundamental para a Pesquisa de Satisfação:

- Na abordagem inicial dos usuários, é importante explicar o que se quer fazer com a pesquisa, o seu objetivo. Além disso, ressaltar a importância da participação de cada um e das informações prestadas para a melhoria do serviço.
- Garantir o caráter confidencial das informações prestadas. É necessário explicar que as informações não serão usadas individualmente, mas apenas em seu conjunto, para estabelecer padrões gerais.
- É importante informar o tempo de entrevista (no máximo 5 minutos), inclusive como elemento de convencimento dos usuários.
- Evitar a solicitude de ajudar os entrevistados a responder o questionário. Esse tipo de comportamento pode induzir, mesmo que não intencionalmente, as respostas de seu entrevistado.
- O entrevistador não deve dar “dicas” para o entrevistado, opinar ou discutir as opiniões dos seus respondentes. Quando um entrevistado demonstra dúvida ou má compreensão, o entrevistador deve simplesmente repetir a pergunta integralmente.
- Na medida do possível, devem-se padronizar os procedimentos da entrevista, desde a seleção do entrevistado até a forma de abordar o usuário e a forma de fazer as perguntas.

IV. Metodologia de análise

Com relação à amostra, seguem algumas ponderações:

Amostragem é o estudo de um pequeno grupo de elementos retirado de uma população que se pretende conhecer. Trata-se de uma técnica de pesquisa na qual um sistema preestabelecido de amostras é considerado idôneo para representar o universo pesquisado, com margem de erro aceitável.

Há vários métodos de amostragem e, neste trabalho, optou-se pela amostragem por proporções, pois a variância é desconhecida. A determinação do tamanho da amostra para populações finitas é feita, na amostragem por proporção, por meio da seguinte fórmula:

$$n = \frac{Npqz^2}{(N - 1) E^2 + pqz^2}$$

Onde:

n é o tamanho da amostra;

N é o tamanho da população;

p é a probabilidade de sucesso;

q é a possibilidade de insucesso;

z é a abscissa da distribuição normal padrão;

E é o erro; e

$p = 1 - q$.

Como não se conhece os valores de p (proporção populacional de indivíduos que pertence à categoria que interessa estudar) e q (proporção populacional de indivíduos que NÃO pertence à categoria que interessa estudar), utilizou-se $p=q= 0,5$ que é a variabilidade máxima. Trabalha-se com 95% de confiança, isto é, se fizermos a

pesquisa 100 vezes, em 95 encontraríamos o mesmo resultado. O erro aceitável foi de 5%.

Para assegurar um menor erro amostral, é muito importante que a amostra seja aleatória. Os entrevistadores deverão ser treinados para minimizar esse erro. Amostragem aleatória é uma técnica que visa a assegurar que a probabilidade de escolha de cada um seja igual à probabilidade de seleção de qualquer outro na mesma população, de modo que todos – e qualquer um dos componentes da população – tenham a mesma probabilidade de participar da pesquisa.

Considerando essas ponderações, chegou-se ao seguinte resultado quanto ao cálculo da amostra:

Hospital Universitário	População		Tamanho da Amostra	
	Consultas	Internações	Consultas	Internações
CHC – UFPR	470.357	15.529	380	296
HC – UFG	228.252	10.248	376	265
HC - UFMG	291.648	15.804	378	297
HC – UFPE	223.597	12.249	383	372
HCTM – UFTM*	157.229	17.887	373	305
HRL/UFS	3.720	53.124	172	353
HU – UFJF	77.716	2.612	362	139
HU - UFMA	356.799	13.466	379	286
HU - UFPEL		5.352		206
HU – UFPI	90.194	3.565	365	167
HU – UFS	112.080	2.364	369	130
HU – UFSM	162.588	13.217	373	285
HU - UNIVASF	15.888	6.384	298	223
HUAB - UFRN	50.073	3.037	352	153
HUB – UNB	151.320	8.304	373	247
HUCAM - UFES	169.056	9.648	374	260
HU-FURG	116.447	6.405	369	223
HUGD – UFGD	41.132	9.661	345	260

HUGV - UFAM	117.133	5.529	369	210
HUJB- UFPB	27.414	409	329	32
HUJM - UFMT	112.676	2.585	369	138
HULW - UFPB	93.192	6.864	367	188
HUMAP - UFMS	97.588	8.152	367	245
HUOL – UFRN	218.687	7.897	376	243
HUPAA - UFAL	110.028	7.824	369	242
HUPES - UFBA	228.432	5.664	376	212
HU-UFSCAR	101.616	576	367	43
HUWC – UFC	207.957	6.373	376	223
MCO – UFBA	111.048	5.460	369	208
MEAC – UFC	49.308	12.012	351	278
MEJC – UFRN	51.976	8.619	353	250

*Dados de 2014

A metodologia a ser utilizada para a mensuração e análise dos dados está descrita no Anexo II.

IV. Cronograma

ETAPA	PRAZO
ABRIL	
Publicação do material de divulgação	A partir de 04/04
MAIO / JUNHO	
Início da aplicação do questionário	02/05
Término da aplicação do questionário	31/05
Geração dos resultados e elaboração do Relatório	06/06

ANEXO I. Questionário da Pesquisa de Satisfação

Olá,

Sou pesquisador do Hospital [nome do hospital] e gostaria de contar com a atenção do(a) senhor(a), por alguns segundos, para responder rapidamente seis perguntas simples sobre a sua opinião em relação ao atendimento do hospital.

Peço, por favor, que o(a) senhor(a) sinta-se bem à vontade para dar a sua opinião sincera. A pesquisa é anônima e sua opinião vai nos ajudar muito na melhoria do hospital.

Vamos lá?

Registrar o sexo do respondente: M () F ()

1. Qual é a sua idade? _____

2. Que serviço o(a) senhor(a) utilizou hoje, aqui no Hospital?

3. Agora, vou fazer algumas perguntas para saber a sua opinião sobre a estrutura do Hospital.

Diga como se sentiu em relação ao seguinte:

ITENS	ÓTIMO	BOM	REGULAR	RUIM	PÉSSIMO
3.1. Conforto no local da recepção <i>(exemplos: assentos, bebedouros e banheiros)</i>					
3.2. Higiene, limpeza e organização do Hospital					
3.3. Conforto das instalações na área de atendimento médico <i>(exemplos: consultórios e locais de espera)</i>					

4. Agora, vou querer saber a sua opinião sobre o atendimento que recebeu.

Diga como se sentiu em relação ao seguinte:

ITENS	ÓTIMO	BOM	REGULAR	RUIM	PÉSSIMO
4.1. Atendimento da recepção: gentileza, atenção, informações recebidas					
4.2. Atendimento da equipe de saúde: gentileza e tratamento de saúde recebido					
4.3. Tempo de espera pelo atendimento/internação					

5. Por fim, quero saber de um modo geral, como o(a) senhor(a) se sentiu quanto ao atendimento recebido no Hospital?

OBS: forma de perguntar:

Satisfeito ou insatisfeito?

Apenas [(in)satisfeito] ou muito [(in)satisfeito]?

ITEM	MUITO INSATISFEITO	INSATISFEITO	SATISFEITO	MUITO SATISFEITO
5.1. Atendimento geral				

6. O(A) Senhor(a) indicaria esse Hospital para algum familiar seu?

[] sim [] não

7. O(A) senhor(a) estudou?

Se não: sabe ler e escrever?

Se sim: até que série?

Registrar a escolaridade:

- () Não sabe ler/escrever
- () Alfabetizado (pessoa capaz de ler um bilhete simples)
- () Nível fundamental completo (1º grau)
- () Nível médio completo (2º grau)
- () Superior completo
- () Pós-graduação

Rodrigo Peñaloza, DAF/EBSERH

Agosto de 2013

1. Introdução

Em geral, qualidade é definida como a satisfação das expectativas do consumidor. No setor de saúde a definição de consumidor e os critérios de qualidade são elementos mais complicados do que os seus correspondentes nas demais indústrias do mercado. Evidentemente, o consumidor é o paciente que recebe o tratamento médico, mas o que ele recebe não é algo que ele compreende inteiramente. Além disso, a concepção de qualidade do hospital difere entre os três agentes envolvidos: (a) o paciente propriamente dito, (b) os médicos e demais servidores do hospital e (c) os agentes provedores, como o governo e as seguradoras. A definição mais aceita internacionalmente é a do *Institute of Medicine* (1990, vol. 2, cap. 5, pp. 128-129)¹, segundo o qual “qualidade da atenção à saúde é o grau em que os serviços de saúde para indivíduos e populações aumentam a probabilidade dos resultados desejados, desde que compatíveis com o atual conhecimento profissional”.

Neste documento apresentamos um método de avaliação da qualidade do hospital a partir do julgamento do usuário, o paciente. O primeiro método sugerido é o da escalagem de Lickert², uma técnica psicométrica que permite transformar em notas numéricas os julgamentos categóricos dos avaliadores, obtendo, assim, uma nota para o hospital. Em seguida oferecemos dois métodos mais simples: o índice de balanço e o índice de qualidade. O índice que sugerimos seja usado na avaliação de qualidade por parte do usuário-paciente é o índice de qualidade com quatro ou cinco categorias ordenadas (m=4 ou 5).

¹ IOM (Institute of Medicine). *Medicare: A Strategy for Quality Assurance*. K.N. Lohr, ed. Washington, D.C.: National Academy Press, 1990.

² Likert, Rensis: "A technique for the measurement of attitudes". *Archives of Psychology*, **140**: 1–55 (1932).

2. O método da escalagem de Lickert

No “Relatório preliminar da pesquisa de satisfação com cidadãos usuários e não usuários do SUS quanto aos aspectos de acesso e qualidade percebida na atenção à saúde e/ou urgência e emergência, mediante inquérito amostral”, do Ministério da Saúde (2011), encontra-se um questionário com 35 perguntas, além de perguntas específicas sobre o perfil do respondente. Algumas dessas perguntas têm caráter nominal e outras, a maioria, têm caráter ordinal. Um exemplo de pergunta de caráter nominal é a pergunta 1, “Para qual(is) atendimento(s) você procurou o SUS (pode marcar mais de uma opção)”, para a qual as possíveis respostas são: vacinação, curativos, orientações; consultas médicas; consulta odontológica; exames de laboratório ou imagem; atendimentos de urgência/SAMU; internações e, finalmente, para pegar medicamentos. Um exemplo de pergunta de caráter ordinal é a pergunta 9, “Como você avalia o atendimento da equipe de enfermagem?”, para a qual as possíveis respostas são: muito ruim, ruim, regular, bom e muito bom.

A qualidade de um hospital é um construto cujos diversos aspectos podem ser associados a itens a cuja mensuração se deve proceder. Em um questionário, cada um desses diversos itens nada mais é que um pergunta e a cada uma delas deve ser atribuída uma nota. O questionário citado visa evidentemente a avaliar a qualidade do SUS. Dizer que essa qualidade é um construto é dizer que o que se entende por qualidade é algo que se constrói. O conjunto das perguntas é, assim, a coleção dos aspectos do construto. Um aspecto é a atenção da equipe de enfermagem, outro, a dos médicos, etc. Quando a pergunta que caracteriza um aspecto é uma variável ordinal categórica, o respondente não atribui uma nota numérica, mas escolhe uma categoria que, em seu julgamento, melhor caracteriza sua percepção sobre o aspecto avaliado. Os métodos de escalagem psicométrica procuram, com base em pressupostos estatísticos, transformar um julgamento categórico em uma nota numérica.

A escalagem de Lickert é um método psicométrico usado em questionários compostos de asserções em relação às quais o respondente indica o seu grau de concordância ou discordância de acordo com uma escala ordinal categórica. Com base nas respostas, procede-se a um mecanismo estatístico de extração de notas numéricas latentes dos

respondentes para cada asserção, obtendo-se, assim, a nota média do hospital. Para ilustrarmos o que é a escalagem de Lickert em termos mais claros, basta que reescrevamos a pergunta 9 do Relatório Preliminar acima como uma asserção da seguinte forma: “o atendimento da equipe de enfermagem foi bom”, oferecendo ao respondente as seguintes alternativas: discordo fortemente, discordo, indeciso, concordo, concordo fortemente. Note que essas cinco categorias são ordenadas, desde uma avaliação fortemente negativa até uma fortemente positiva.

Note que a asserção acima possui um tom positivo, pois assevera que o atendimento da equipe de enfermagem foi bom, cabendo ao respondente apenas expressar o seu grau de discordância ou concordância. Uma asserção também pode ter um tom negativo, cabendo igualmente ao respondente expressar seu grau de discordância ou concordância. Por exemplo, no *Long-Form Patient Satisfaction Questionnaire (PSQ-III)*, RAND Corporation, a pergunta 7 é a asserção negativa “geralmente eu fico esperando por longo tempo no consultório médico”, sobre cujo grau de concordância o respondente deve se manifestar.

No método de Lickert, deve-se ter um grande número de itens, cada item sendo uma asserção, metade das quais positivas, a outra metade negativa, distribuídas aleatoriamente. O objetivo dessa condição é não induzir o respondente a um viés positivo ou negativo.

2.1. Cálculo dos scores individuais

Considere um questionário $Q = \{A_1, \dots, A_N\}$ com um grande número de itens, A_1, \dots, A_N . Cada item é uma asserção e pode ter ou um tom positivo ou um tom negativo relativamente à concepção que se faz do aspecto do construto. Sem perda de generalidade, suponha que N é par e que $A_1, \dots, A_{N/2}$ são asserções positivas e que $A_{(N/2)+1}, \dots, A_N$ são asserções negativas. No questionário real aplicado aos respondentes, a distribuição dessas asserções deve ser aleatória. Para a escalagem de Lickert propriamente dita, isso é irrelevante. Se a asserção é positiva, o respondente deve escolher uma, e apenas uma, das seguintes categorias ordenadas: 1 – discordo fortemente, 2 – discordo, 3 – indeciso, 4 – concordo e 5 – concordo fortemente. Se a

asserção é negativa, a ordenação apresentada das categorias se mantém, mas a numeração se inverte: 5 – discordo fortemente, 4 – discordo, 3 – indeciso, 2 – concordo e 1 – concordo fortemente. Para cada asserção A_n , $n = 1, \dots, N$, computa-se a frequência relativa f_{nk} de cada categoria k :

categoria	frequência relativa
1	f_{n1}
2	f_{n2}
3	f_{n3}
4	f_{n4}
5	f_{n5}
total	1

Observe que a representação acima comporta tanto asserções positivas quanto negativas. Com efeito, qualquer que seja o tom do item, positivo ou negativo, as alternativas de resposta devem ser sempre ordenadas desde “discordo fortemente” até “concordo fortemente”, seja da esquerda para a direita, seja de cima para baixo. Quando a asserção é positiva, às alternativas disponíveis associam-se os números 1 a 5, nessa ordem, pois assim denota-se um melhoramento. Quando a asserção é negativa, associam-se às alternativas disponíveis os números de 5 a 1 de forma respectiva. Por exemplo, se numa asserção positiva, o respondente escolhe a categoria “concordo fortemente”, essa escolha corresponde ao número 5, que é o melhor julgamento. Numa asserção negativa, se ele escolhe “discordo fortemente”, ele está, na verdade, dando a melhor avaliação possível àquele item, razão porque a essa escolha corresponde o número 5. Dessa forma, independentemente do tom da asserção, se positivo ou negativo, o número 1 sempre corresponderá à avaliação mais negativa do hospital no que tange àquele item e o número 5 corresponderá à mais positiva.

Um pressuposto importante da escalagem psicométrica é que o respondente, ao avaliar um item, atribui um valor quantitativo a esse item, valor esse expresso por um número real. Desse pressuposto deriva o que Thurstone denominou de *continuum psicológico*. Os respondentes, do ponto de vista amostral, teriam critérios de julgamento diferentes, mas de tal sorte que a avaliação média e sua discrepância (dada pela variância) seriam

parâmetros de uma distribuição normal. Em outras palavras, pressupõe-se que os julgamentos numéricos são normalmente distribuídos. Para uma amostra grande de respondentes, essa suposição seria razoável. O problema é que não observamos esses valores de escala quantitativos, mas apenas as suas manifestações na escala categórica ordinal, sendo aqueles valores de escala, portanto, *variáveis latentes*. Lickert admitiu, na mesma linha de Thurstone, que os valores de escala que, no *continuum* psicológico, representam as avaliações quantitativas que vão se expressar, como atitudes, na escala categórica ordinal, têm distribuição normal. O processo de categorização provocado pela natureza ordinal do item avaliado equivale a substituir a curva normal contínua por um histograma de cinco classes.

Seja $F(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ a função de distribuição de probabilidade da normal-padrão $\mathcal{N}(0,1)$. Se uma percentagem f_{n1} manifesta uma opinião maximamente negativa do item avaliado, considere, inicialmente, o valor crítico θ_{n1} tal que $F(\theta_{n1}) = f_{n1}$, ou seja, o valor crítico para o qual a percentagem dos que fazem a avaliação mais baixa é igual à área da cauda esquerda da curva normal-padrão. Essa área representa a proporção dos que escolheram a categoria inferior. O valor de escala atribuído a essa primeira categoria é o valor crítico correspondente à metade dessa área, a saber, o valor crítico z_{n1} tal que $F(z_{n1}) = f_{n1}$, isto é, z_{n1} tal que $\int_{-\infty}^{z_{n1}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{f_{n1}}{2}$. Em outras palavras, é o ponto médio do intervalo de classe correspondente à categoria inferior. Para a segunda categoria, posterior à inferior, consideramos o ponto médio do intervalo para o qual a área sob a curva normal é igual à proporção f_{n2} , a saber, o valor crítico z_{n2} tal que $F(z_{n2}) = f_{n1} + \frac{f_{n2}}{2}$, isto é, z_{n2} tal que $\int_{-\infty}^{z_{n2}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = f_{n1} + \frac{f_{n2}}{2}$. Prosseguindo dessa forma, temos, então, os seguintes valores de escala para as categorias:

$$z_{n1} = F^{-1}\left(\frac{f_{n1}}{2}\right)$$

$$z_{n2} = F^{-1}\left(f_{n1} + \frac{f_{n2}}{2}\right)$$

$$z_{n3} = F^{-1}\left(f_{n1} + f_{n2} + \frac{f_{n3}}{2}\right)$$

$$z_{n4} = F^{-1}\left(f_{n1} + f_{n2} + f_{n3} + \frac{f_{n4}}{2}\right)$$

$$z_{n5} = F^{-1}\left(f_{n1} + f_{n2} + f_{n3} + f_{n4} + \frac{f_{n5}}{2}\right)$$

Assim, um indivíduo que respondeu ao item A_n escolhendo a categoria k , em que $k = 1, \dots, 5$, terá nota z_{nk} . Suponha que I indivíduos são amostrados. Denote por $k(n, i)$ a categoria escolhida pelo indivíduo $i = 1, \dots, I$ com respeito ao item A_n , $n = 1, \dots, N$. A nota do indivíduo i será a soma dos *scores* correspondentes a cada item, ou seja, o *score* total do indivíduo i é dado por $s_i = \sum_{n=1}^N z_{n,k(n,i)}$. Se não houver qualquer seleção dos itens com base em critérios estatísticos de discriminação, então a nota final do hospital $h = 1, \dots, H$ será a média dos seus *scores* individuais, $\bar{s} = \frac{1}{I} \sum_{i=1}^I s_i$. Entretanto, a etapa seguinte consiste exatamente na seleção dos itens a serem considerados. Para simplificar, defina:

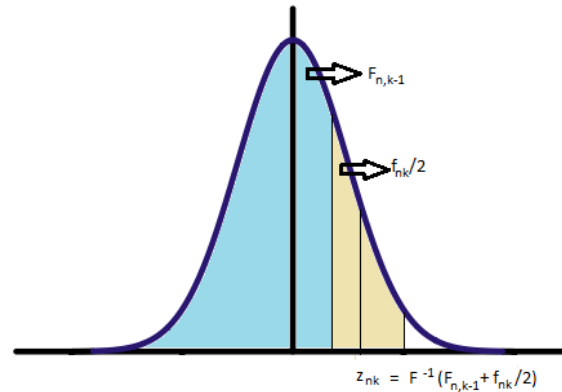
$$\begin{aligned}\Phi_{n1} &= \frac{f_{n1}}{2} \\ \Phi_{n2} &= f_{n1} + \frac{f_{n2}}{2} \\ \Phi_{n3} &= f_{n1} + f_{n2} + \frac{f_{n3}}{2} \\ \Phi_{n4} &= f_{n1} + f_{n2} + f_{n3} + \frac{f_{n4}}{2} \\ \Phi_{n5} &= f_{n1} + f_{n2} + f_{n3} + f_{n4} + \frac{f_{n5}}{2}\end{aligned}$$

Assim, $z_{nk} = F^{-1}(\Phi_{nk})$, para $k = 1, \dots, 5$. Observe que, se $F_{nk} = f_{n1} + f_{n2} + \dots + f_{nk}$ denota a frequência acumulada até a k^{a} categoria, então $\Phi_{nk} = F_{n,k-1} + \frac{f_{nk}}{2}$, o que mostra que Φ_{nk} é uma espécie de valor intermédio entre $F_{n,k-1}$ e F_{nk} . Por isso, vamos chamar Φ_{nk} de *frequência proto-acumulada*, no sentido de ser anterior à frequência acumulada propriamente dita. Assim, o valor de escala da categoria k do item n é o valor numérico z_{nk} para o qual:

$$\begin{aligned}\int_{-\infty}^{z_{nk}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx &= \Phi_{nk} \\ &= F_{n,k-1} + \frac{f_{nk}}{2} \\ &= f_{n1} + \dots + f_{n,k-1} + \frac{f_{nk}}{2}\end{aligned}$$

O gráfico seguinte ilustra como determinar z_{nk} mediante uma tabela de valores críticos da distribuição normal-padrão.

Gráfico 1: valor de escala na curva normal-padrão.



Dado o valor Φ_{nk} , calculamos o valor de escala z_{nk} para o qual a área sob a curva da normal-padrão à esquerda de z_{nk} é exatamente $\Phi_{nk} = F_{n,k-1} + \frac{f_{nk}}{2}$.

Imagine o seguinte caso extremo segundo o qual $f_{n1} = 1$ e $f_{n2} = \dots = f_{n5} = 0$. Então, $z_{n1} = F^{-1}(0,5)$, donde $z_{n1} = 0$. Além disso, $z_{n2} = \dots = z_{n5} = F^{-1}(1)$, donde $z_{n2} = \dots = z_{n5} = +\infty$. Imagine o outro caso extremo segundo o qual $f_{n5} = 1$ e $f_{n1} = \dots = f_{n4} = 0$. Então, $z_{n1} = \dots = z_{n4} = F^{-1}(0)$, donde $z_{n1} = \dots = z_{n4} = -\infty$. Além disso, $z_{n5} = F^{-1}(0,5)$, ou seja, $z_{n5} = 0$. Entre esses dois extremos, considere o caso de uma distribuição concentrada, digamos, na segunda categoria, posterior à inferior. Nesse caso, $z_{n1} = -\infty$, $z_{n2} = 0$ e $z_{n3} = z_{n4} = z_{n5} = +\infty$. Algo similar acontece com qualquer distribuição concentrada. Nesses casos os *scores* não estão bem definidos e não há como somá-los. Porém, como veremos adiante, itens que apresentem respostas concentradas em uma categoria não serão selecionadas, pois não passarão no teste-*t* de diferença de médias, pois eles não discriminariam os respondentes em relação às suas avaliações. Sempre que a categoria inferior e/ou a superior tiverem frequência relativa nula, o item será rejeitado pelo teste-*t*.

2.2. Seleção de itens

Dado um item $n = 1, \dots, N$ qualquer, considere a ordenação ascendente de seus respondentes de acordo com os *scores* individuais: $z_{n,k(n,i_1)} \leq z_{n,k(n,i_2)} \leq \dots \leq z_{n,k(n,i_l)}$, ou seja, as estatísticas de ordem dos *scores* para o item n . Considere ainda o subgrupo de respondentes formado pelo 1º quartil, isto é, $I_\alpha = \{i_1, \dots, i_{\lfloor N/4 \rfloor}\}$, e seja $\alpha = \{z_{n,k(n,i_1)} \leq \dots \leq z_{n,k(n,i_{\lfloor N/4 \rfloor})}\}$ a lista ordenada de seus *scores* para o item n . Analogamente, considere o subgrupo formado pelo 4º quartil, isto é, $I_\omega = \{i_{\lfloor 3N/4 \rfloor + 1}, \dots, i_N\}$, e seja $\omega = \{z_{n,k(n,i_{\lfloor 3N/4 \rfloor + 1})} \leq \dots \leq z_{n,k(n,i_N)}\}$ a lista ordenada de seus *scores* para o item n , em que $\lfloor x \rfloor$ denota a parte inteira de um número. Defina as seguintes estatísticas:

$$\mu_n(\alpha) = \frac{1}{N_\alpha} \sum_{\ell \in I_\alpha} z_{n,k(n,\ell)}: \text{média dos } \textit{scores} \text{ do grupo inferior, } \alpha$$

$$\sigma_n^2(\alpha) = \frac{1}{N_\alpha - 1} \sum_{\ell \in I_\alpha} (z_{n,k(n,\ell)} - \mu_n(\alpha))^2: \text{variância dos } \textit{scores} \text{ do grupo inferior, } \alpha$$

$$N_\alpha = \lfloor N/4 \rfloor: \text{número de respondentes no grupo inferior, } \alpha$$

$$\mu_n(\omega) = \frac{1}{N_\omega} \sum_{\ell \in I_\omega} z_{n,k(n,\ell)}: \text{média dos } \textit{scores} \text{ do grupo superior, } \omega$$

$$\sigma_n^2(\omega) = \frac{1}{N_\omega - 1} \sum_{\ell \in I_\omega} (z_{n,k(n,\ell)} - \mu_n(\omega))^2: \text{variância dos } \textit{scores} \text{ do grupo superior, } \omega$$

$$N_\omega = N - \lfloor 3N/4 \rfloor: \text{número de respondentes no grupo inferior, } \omega$$

Na avaliação das respostas dos grupos inferior e superior, consideramos o teste t da diferença entre duas médias, dado pela estatística:

$$t = \frac{\mu_n(\omega) - \mu_n(\alpha)}{\sqrt{\sigma_n^2(\alpha)/N_\alpha + \sigma_n^2(\omega)/N_\omega}}$$

Quando $N_\alpha = N_\omega = \nu$, como geralmente será o caso, a fórmula acima se reduz a:

$$t = \sqrt{\nu} \frac{\mu_n(\omega) - \mu_n(\alpha)}{\sqrt{\sigma_n^2(\alpha) + \sigma_n^2(\omega)}}$$

O valor de t é uma medida da extensão em que a asserção n discrimina entre os grupos inferior e superior. Quanto maior o valor de t , mais a asserção discrimina entre os dois subgrupos. Como regra de bolso, ainda que crua e aproximada, podemos interpretar qualquer valor $t \geq 1,75$ como indicando que as respostas médias dos grupos inferior e superior diferem significativamente, desde que os grupos em questão tenham pelo menos 25% dos respondentes cada um, o que é o caso limite quando tomamos o 1º e o

4º quartil. Assim, se t_n denota o valor- t do item n , selecione somente aqueles itens para os quais $t_n \geq 1,75$. Denote por $\mathfrak{A} = \{A_n \in \mathcal{Q}: t_n \geq 1,75\}$ o conjunto de itens selecionados pelo teste- t . Por fim, denote por $\mathfrak{S} = \{n: A_n \in \mathfrak{A}\}$ o conjunto dos índices dos itens selecionados. Usualmente, o número de itens é de 40 a 50, mas recomendamos que o questionário de avaliação dos hospitais não seja muito grande, de modo a evitar transtornos aos usuários.

2.3. Notas dos hospitais

A nota do indivíduo i atribuída ao hospital será a soma dos *scores* correspondentes aos itens selecionados, ou seja, o *score* do indivíduo i é dado por $S_i(h) = \sum_{n \in \mathfrak{S}} Z_{n,k(n,i)}$. Observe que indexamos o *score* individual de acordo com o hospital avaliado, pois, para cada hospital, procedemos às etapas descritas acima. A nota final do hospital $h = 1, \dots, H$ será a média dos seus *scores* individuais:

$$\bar{S}(h) = \frac{1}{I(h)} \sum_{i=1}^{I(h)} S_i(h)$$

em que $I(h)$ é o número de respondentes do hospital h . Quando da tabulação dos dados, é importante identificar as categorias escolhidas por cada respondente em cada item do questionário, de forma a obter-se a nota final por ele atribuída. As notas não devem ser comparadas entre hospitais. Como as populações das quais são tomadas as amostras não são as mesmas e como os itens selecionados em um hospital não necessariamente coincidem os selecionados em outros, as notas devem ser usadas pelo hospital como um instrumento de avaliação próprio, como um *feedback* de seus usuários, indicando-lhe o estado dos seus serviços em termos de qualidade.³

2.4. Exemplos

³ O tamanho da amostra de respondentes pode ser calculado mediante a fórmula $S = \frac{z_c^2 \pi(1-\pi)}{D^2}$, em que S é o tamanho da amostra, z_c é o valor crítico associado a um determinado nível de confiança para a distribuição normal-padrão, π é o percentual esperado do evento para a população e D é o erro amostral. Em seguida, ajusta-se pelo tamanho da população, P , de acordo com a relação $S^* = \frac{S}{1+S/P}$.

Suponha que 1.000 indivíduos responderam a um questionário e que as respostas a um dos itens, digamos, n , tenham apresentado as frequências absolutas conforme a tabela abaixo:

item n				
categoria	freq. abs.	freq. rel. f_{n1}	freq. proto-ac. Φ_{nk}	valor de escala $z_{nk} = F^{-1}(\Phi_{nk})$
1	100	$f_{n1} = 0,10$	$\Phi_{n1} = 0,050$	-1,64
2	180	$f_{n2} = 0,18$	$\Phi_{n2} = 0,190$	-0,88
3	220	$f_{n3} = 0,22$	$\Phi_{n3} = 0,390$	-0,28
4	350	$f_{n4} = 0,35$	$\Phi_{n4} = 0,675$	0,46
5	150	$f_{n5} = 0,15$	$\Phi_{n5} = 0,925$	1,42
total	1000	1		

O 4º quartil é composto de 150 respondentes com *score* 1,42 e 100 com *score* 0,46. O *score* médio desse grupo é $\mu_n(\omega) = \frac{150 \times (1,42) + 100 \times (0,46)}{250} = 1,036$ e a variância é $\sigma_n^2(\omega) = \frac{100 \times (1,42 - 1,036)^2 + 150 \times (0,46 - 1,036)^2}{249} = 0,2591$. Já no 1º quartil de respondentes, os 25% com os menores *scores*, teremos 100 respondentes com *score* -1,64 e 150 com *score* -0,88. O *score* médio desse grupo é $\mu_n(\alpha) = -1,184$ e a variância é $\sigma_n^2(\alpha) = 0,1392$. O valor- t é $t = \sqrt{250} \frac{1,036 - (-1,184)}{\sqrt{0,1392 + 0,2591}}$, ou seja, $t = 55,6183$. O item é, por conseguinte, selecionado.

2.5. As planilhas

A parte inicial do trabalho pode ser feita numa planilha Excel bastante simples, como ilustrada a seguir. Os indivíduos são listados ao longo das linhas e os itens com as suas respectivas categorias ao longo das colunas. Cada indivíduo, em cada item, escolhe uma única alternativa, em cuja respectiva entrada deve-se inserir o valor numérico 1. Em todas as demais, no item, insere-se 0 (zero). Abaixo das linhas dos indivíduos há uma linha com o total de marcações de cada alternativa em cada item. Cada um desses totais é dividido pelo número de respondentes ou indivíduos a fim de se obter as frequências relativas de cada alternativa dentro de cada item. Numa outra linha inserimos as frequências proto-acumuladas. Por fim, uma linha com os respectivos valores de escala, obtidos pela inversa da função de distribuição da normal-padrão, disponível na própria

planilha Excel. Tudo que expusemos acima refere-se ao item, devendo-se repetir todo o procedimento descrito item a item.

indivíduo	item 1					item 2					item N				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
2	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0
3	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
I	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0
total	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{14}	Γ_{15}	Γ_{21}	Γ_{22}	Γ_{23}	Γ_{24}	Γ_{25}	Γ_{N1}	Γ_{N2}	Γ_{N3}	Γ_{N4}	Γ_{N5}
freq. relativa $f_{nk} = \frac{\Gamma_{nk}}{I}$	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{21}	f_{22}	f_{23}	f_{24}	f_{25}	f_{N1}	f_{N2}	f_{N3}	f_{N4}	f_{N5}
freq. proto-ac. $\Phi_{nk} = \frac{\sum_{\ell=1}^{k-1} f_{n\ell} + \frac{f_{nk}}{2}}$	Φ_{11}	Φ_{12}	Φ_{13}	Φ_{14}	Φ_{15}	Φ_{21}	Φ_{22}	Φ_{23}	Φ_{24}	Φ_{25}	Φ_{N1}	Φ_{N2}	Φ_{N3}	Φ_{N4}	Φ_{N5}
Valor de escala $Z_{nk} = F^{-1}(\Phi_{nk})$	Z_{11}	Z_{12}	Z_{13}	Z_{14}	Z_{15}	Z_{21}	Z_{22}	Z_{23}	Z_{24}	Z_{25}	Z_{N1}	Z_{N2}	Z_{N3}	Z_{N4}	Z_{N5}

Uma vez determinados os valores de escala, construímos outras planilhas, uma para cada item, substituindo as categorias escolhidas, no item, pelos respectivos valores de escala. Em outras palavras, no item 1 acima substituímos, na linha correspondente ao indivíduo 1, o valor numérico 1 na 4ª categoria por Z_{14} ; na linha correspondente ao indivíduo 2, substituímos o valor numérico 1 na 3ª categoria por Z_{13} e assim por diante. Abaixo ilustramos com uma planilha para o item 1:

indivíduo	item 1				
	1	2	3	4	5
1	0	0	0	Z_{14}	0
2	0	0	Z_{13}	0	0
3	0	0	0	Z_{14}	0
⋮	⋮	⋮	⋮	⋮	⋮
I	0	0	0	Z_{14}	0

Em seguida, ordenamos as linhas de acordo com os valores de escala, do menor ao maior. Calculamos a média e variância do grupo correspondente ao 1º quartil e do grupo correspondente ao 4º quartil, para enfim aplicar o teste-t. se o item não passar no teste,

ele é excluído. Se a nota média do hospital é negativa, então é porque os usuários não concordam com a asserção de que a qualidade do hospital é boa. Quanto mais negativa, mais eles discordam da qualidade do hospital. Se for positiva, concordam. Quanto maior o número, maior o grau de concordância.

O método da escalagem de Lickert requer que o respondente mantenha um grau de discernimento e atenção razoavelmente elevado durante o preenchimento do questionário. De fato, independentemente do seu grau intelectual, a obrigação de responder a uma série longa de decisões caracterizadas pela escolha de uma dentre cinco categorias ordenadas pode ser mentalmente cansativa, ainda mais se se considerar que o preenchimento se dá quando o usuário do HU acabou de passar por um momento de sofrimento físico e psicológico. É importante, portanto, que apresentemos metodologias alternativas mais simples e que não se constituam em um peso adicional para o usuário.

3. Índice de balanço

Considere uma vez mais um questionário $Q = \{Q_1, \dots, Q_N\}$, em que cada Q_i , para $i = 1, \dots, N$, é uma pergunta, não mais uma asserção. A pergunta Q_i é uma variável ordinal, ou seja, as possíveis respostas são dadas por uma sequência de categorias ordenadas da pior para a melhor. Para o índice de balanço, o número de possíveis categorias é sempre três: uma ruim, uma neutra e uma boa. Por exemplo:

- Q_i : Como foi a atenção que lhe foi dispensada pelo médico no período de internação?
- (1) péssima
 - (2) normal
 - (3) excelente

Observe que apenas três alternativas estão disponíveis. A primeira denota uma avaliação ruim, a segunda uma avaliação neutra e a terceira boa. As perguntas devem ser formuladas de modo tal que as três possíveis categorias sejam ordenadas desde aquela que denote a pior avaliação até aquela que denote a melhor, sendo a intermédica uma categoria que represente uma avaliação neutra ou indiferente ou “normal”, qualquer que seja a concepção que o respondente faça do que seja normal.

O índice de balanço é uma espécie de índice de sentimento. A recomendação da OCDE para índices de sentimento é que o número de perguntas seja pequeno, não mais que 12 ou algo próximo. Recomendamos que para a Pesquisa de Satisfação do Usuário o número de perguntas seja no máximo a 10.

Além dessa recomendação, também propomos algumas características importantes que têm que ser inseridas no questionário. Em primeiro lugar, duas perguntas distintas não devem tratar de temas iguais ou muito relacionados entre si. Deve haver uma espécie de independência semântica entre os temas das perguntas. Em segundo, a totalidade dos temas abordados deve perfazer a integridade do construto inicial que se faça da “qualidade do HU”. Com efeito, a qualidade do HU é um construto que deverá ser definido e esse construto será composto por vários aspectos diferentes entre si. Cada aspecto deve ser específico o suficiente para que tenha significado para o respondente, mas, por outro lado, deve ser geral o suficiente para que todos se sintam capazes de decidir sobre a qualidade do item avaliado. Se uma pergunta for muito específica, ganha-se precisão, mas corre-se o risco de que muitos usuários não se sintam aptos a respondê-la, por não terem vivenciado a experiência caracterizadora da pergunta. Por exemplo, “como você avalia o cuidado da equipe de enfermagem durante os momentos em que você foi submetido(a) a injeção intravenosa?” Essa é uma pergunta demasiadamente específica, pois nem todo usuário terá sido submetido a esse tipo de procedimento. Se a pergunta for muito genérica, todos se sentirão aptos a respondê-la, mas a avaliação final não terá significado. Por exemplo, “Como você avalia a atenção que lhe foi dispensada pela equipe médica, de enfermagem e de demais servidores do HU desde a sua entrada até a sua saída?” Essa pergunta é muito genérica. Deveria ser subdividida em uma pergunta referente somente à equipe médica, outra somente à de enfermagem e outra, se se julgar necessário, aos demais servidores. Outra subdivisão possível poderia ser por etapas: entrada, estadia e saída.

De qualquer forma, a postura que se deve ter quando da elaboração do questionário pode ser resumida numa recomendação que, embora estranha à primeira vista, incorpora muito bem o cuidado e o bom-senso necessários nesse processo: “deve-se

decompor o construto em um número razoavelmente pequeno de aspectos que sejam razoavelmente independentes entre si, razoavelmente específicos porém razoavelmente gerais”.

Para cada pergunta Q_n , $n = 1, \dots, N$, computa-se a frequência relativa f_{nk} de cada categoria k :

categoria	freq. rel.
1	f_{n1}
2	f_{n2}
3	f_{n3}
total	1

O índice de balanço de Q_n é:

$$\mathcal{B}_n = 100 \times (f_{n3} - f_{n1})$$

O índice de balanço exclui, portanto, a massa relativa às opiniões neutras sobre o item avaliado pela pergunta e pondera a massa relativa das opiniões boas contra as más, exatamente como uma balança que penderá para o lado mais pesado. A normalização por 100 tem o intuito de expressar o resultado em termos de números que vão de -100 a +100, em vez de -1 a +1. Quando todos os respondentes avaliam positivamente o item, ou seja, quando a distribuição das respostas é $(f_{n1}, f_{n2}, f_{n3}) = (0,0,1)$, então $\mathcal{B}_n = 100$. Quando todos avaliam negativamente o item, ou seja, quando a distribuição das respostas é $(f_{n1}, f_{n2}, f_{n3}) = (1,0,0)$, então $\mathcal{B}_n = -100$. Quando, independentemente da massa f_{n2} de avaliações neutras, as avaliações extremas têm mesma massa, $f_{n1} = f_{n3}$, então $\mathcal{B}_n = 0$. A avaliação neutra funciona como a haste da balança e as extremas como os pratos.

Considere, porém, os casos: $(f_{n1}, f_{n2}, f_{n3}) = (30\%, 30\%, 40\%)$ e $(\bar{f}_{n1}, \bar{f}_{n2}, \bar{f}_{n3}) = (10\%, 70\%, 20\%)$. Nesses dois casos, os índices de balanço são iguais, $\mathcal{B}_n = \bar{\mathcal{B}}_n = 10$. Não há, porém, como decidir qual dos dois casos é melhor. Portanto, deve-se estar ciente das limitações hermenêuticas do índice de balanço.

A cada pergunta Q_n corresponderá um valor de B_n , de -100 a +100. Sempre que positivo, as avaliações positivas terão superado as negativas e, quanto maior esse valor, maior terá sido a diferença entre as avaliações positivas e negativas.

4. Índice de qualidade

Considere o questionário $Q = \{Q_1, \dots, Q_N\}$, em que cada Q_n , para $n = 1, \dots, N$, é uma variável ordinal. Cada pergunta Q_n oferece uma lista de possíveis respostas categóricas ordenadas da pior para a melhor. Por exemplo:

- Q_n : Como foi a atenção que lhe foi dispensada pelo médico no período de internação?
- (1) péssima
 - (2) ruim
 - (3) normal
 - (4) boa
 - (5) excelente

O número de categorias ordenadas não precisa ser igual para todas as perguntas. Uma pergunta pode ter três possíveis respostas, outra pode ter quatro e outra cinco. Quando o número de possíveis respostas é ímpar, a categoria intermédia deve ter um caráter de neutralidade relativamente à avaliação do item em questão. Quando o número é par, não existe categoria neutra. Por exemplo, se a pergunta exemplificada acima tivesse quatro possíveis respostas, deveríamos ter algo como:

- Q_n : Como foi a atenção que lhe foi dispensada pelo médico no período de internação?
- (1) péssima
 - (2) normal para ruim
 - (3) normal para bom
 - (4) excelente

A justificativa para o número par de respostas é a constatação de que, em questionários longos e dado o estado de cansaço psicológico do paciente, há uma tendência de que o respondente marque a opção neutra em razão de seu cansaço. Quando a alternativa neutra é retirada, o respondente é forçado a avaliar mais acuradamente a sua percepção quanto ao item avaliado. Como isso requer um esforço mental maior, é importante que o questionário, nesse caso, seja curto.

Seja como for, vamos, para simplificar a exposição do método, supor que há m categorias ordenadas, da pior para melhor, para cada pergunta cada Q_n , para $n = 1, \dots, N$. O vetor de frequências relativas das respostas à pergunta Q_n é (f_{n1}, \dots, f_{nm}) . O indicador associado à pergunta é:

$$\mathcal{H}_n = 100 \times \sum_{k=1}^m \left(\frac{k-1}{m-1} \right) f_{nk}$$

Em particular, se $m = 5$, então $\mathcal{H}_n = 25 \times (f_{n2} + 2f_{n3} + 3f_{n4} + 4f_{n5})$. Se $m = 4$, então, $\mathcal{H}_n = \frac{100}{3} \times (f_{n2} + 2f_{n3} + 3f_{n4})$.

Um caso importante é $m = 3$, pois há uma relação direta com o índice de balanço visto na subseção 5.1. Com efeito, para $m = 3$, temos $\mathcal{H}_n = 50 \times (f_{n2} + 2f_{n3})$. Recorde que $\mathcal{B}_n = 100 \times (f_{n3} - f_{n1})$ e que seu espectro de variação é de -100 a +100. Se quiséssemos reescalar o valor de \mathcal{B}_n para que seu espectro de variação seja de 0 a 100, devemos somar 100 a \mathcal{B}_n e dividir o resultado por 2. Com efeito, $\mathcal{B}_n = -100$ será transformado em $\frac{-100+100}{2} = 0$. Já $\mathcal{B}_n = +100$ será transformado em $\frac{+100+100}{2} = 100$. Os valores intermediários entre -100 e +100 serão transformados linearmente para valores entre 0 e 100. Dessa forma, se definirmos o índice de balanço reescalado como $\mathcal{B}_n^* = \frac{\mathcal{B}_n+100}{2}$, teremos:

$$\begin{aligned} \mathcal{B}_n^* &= \frac{100 \times (f_{n3} - f_{n1}) + 100}{2} \\ &= 50 \times (f_{n3} - \underbrace{f_{n1} + 1}_{f_{n2} + f_{n3}}) \\ &= 50 \times (f_{n2} + 2f_{n3}) \end{aligned}$$

Portanto, nesse caso, $\mathcal{B}_n^* = \mathcal{H}_n$. Logo, para $m = 3$, $\mathcal{H}_n = \frac{\mathcal{B}_n+100}{2}$. Alternativamente, $\mathcal{B}_n = 2\mathcal{H}_n - 100$. Essa relação linear simples permite transformar *ad libitum* o índice de balanço no índice de qualidade e *vice-versa*.

5. O questionário de avaliação de qualidade

O questionário a ser utilizado para a avaliação de qualidade contém as cinco perguntas, divididas em dois grupos. O primeiro grupo refere-se à estrutura do hospital e o segundo ao atendimento do pessoal de atenção à saúde.

Com relação à estrutura do hospital, como você se sentiu em relação ao seguinte:

Q_1 : Conforto no local da recepção (assentos, bebedouros e banheiros)?

- (1) ótimo
- (2) bom
- (3) regular
- (4) ruim
- (5) péssimo

Q_2 : Higiene, limpeza e organização do Hospital?

- (1) ótimo
- (2) bom
- (3) regular
- (4) ruim
- (5) péssimo

Q_3 : Conforto das instalações na área de atendimento médico (exemplos: consultórios e locais de espera)?

- (1) ótimo
- (2) bom
- (3) regular
- (4) ruim
- (5) péssimo

Com relação ao atendimento que recebeu, como você se sentiu em relação ao:

Q_4 : atendimento da recepção (gentileza, atenção, informações recebidas)?

- (1) ótimo
- (2) bom
- (3) regular
- (4) ruim
- (5) péssimo

Q_5 : atendimento da equipe de saúde (gentileza e tratamento de saúde recebido)?

- (1) ótimo
- (2) bom
- (3) regular
- (4) ruim
- (5) péssimo

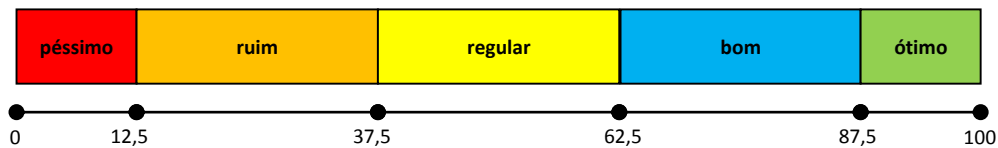
O questionário, portanto, é $\mathcal{Q} = \{Q_1, Q_2, Q_3, Q_4, Q_5\}$, em que cada pergunta possui 5 categorias ordenadas: (1) péssimo, (2) ruim, (3) regular, (4) bom e (5) ótimo. O vetor de frequências relativas das respostas à pergunta Q_n é $(f_{n1}, f_{n2}, f_{n3}, f_{n4}, f_{n5})$. Neste caso, $m = 5$. O indicador associado à pergunta é:

$$\mathcal{H}_n = 25 \times (f_{n2} + 2f_{n3} + 3f_{n4} + 4f_{n5})$$

O indicador agregado do hospital é uma média ponderada dos indicadores \mathcal{H}_n , $n = 1,2,3,4,5$, ou seja, $\bar{\mathcal{H}} = \sum_{n=1}^5 \alpha_n \mathcal{H}_n$, em que $0 < \alpha_n < 1$, para $n = 1,2,3,4,5$, e $\sum_{n=1}^5 \alpha_n = 1$. Existem diversos sistemas de pesos $\{\alpha_n\}$, desde os baseados em Análise Fatorial até os mais simples como o determinado pela média aritmética. Por simplicidade, escolhemos a média aritmética dos indicadores das perguntas como o indicador agregado do hospital. Assim, o índice de qualidade agregado do hospital é:

$$\bar{\mathcal{H}} = \frac{1}{5} \sum_{n=1}^5 \mathcal{H}_n$$

Cumpre esclarecer o significado prático do indicador de qualidade. Dada uma pergunta Q_n , considere o seu índice de qualidade \mathcal{H}_n . Em primeiro lugar, \mathcal{H}_n varia de 0 a 100. O intervalo $[0,100]$ é decomposto em quatro subintervalos, cujos *thresholds* são os indicados no gráfico a seguir:



É possível associar o valor numérico obtido para \mathcal{H}_n a uma das categorias de satisfação. Se, por exemplo, $\mathcal{H}_n = 60$, que é um número entre 37,5 e 62,5, podemos dizer que a opinião “média” dos respondentes é *regular*. Se o valor fosse $\mathcal{H}_n = 75$, podemos dizer que a opinião “média” é *bom*, porém ainda assim podemos efetuar uma comparação do grau de satisfação pela comparação entre os valores 75 e 60. Neste último caso, os usuários estão mais satisfeitos do que no primeiro caso. Temos, portanto, não apenas um indicador numérico, que nada mais é que uma medida de posição para variáveis ordinais, mas também um indicador categórico, ou seja, uma medida de posição categórica para variáveis ordinais.

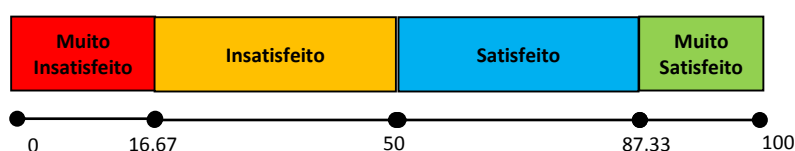
Quando 100% dos respondentes manifestam a opinião *péssimo*, $\mathcal{H}_n = 0$. Mas pode ocorrer que 100% dos respondentes opinem *ótimo*, em cujo caso $\mathcal{H}_n = 100$. Se 50% opinam *ruim* e 50% *bom*, então $\mathcal{H}_n = 50$ e, além disso, podemos também dizer que a opinião “média” é *regular*.

Sempre que \mathcal{H}_n for igual a algum dos *thresholds* indicados acima, há uma indeterminação. Por exemplo, se $\mathcal{H}_n = 62,5$, a medida de posição categórica é o conjunto {regular, bom}.

Uma análise similar pode ser feita para o índice de qualidade agregado. É importante ressaltar, porém, que a essa análise subjaz uma escolha de pesos particular. A razão da escolha da média aritmética é não dar mais importância a um item avaliado do que a qualquer outro.

Se forem acrescentadas mais perguntas, procede-se da mesma forma exposta acima. A única diferença é que a nota final do hospital será uma média aritmética entre um número maior de valores.

Recorde que, se $m = 4$, então, $\mathcal{H}_n = \frac{100}{3} \times (f_{n2} + 2f_{n3} + 3f_{n4})$. Neste caso, não há categoria neutra (regular, indiferente etc.) e os *thresholds* serão os números indicados no desenho abaixo:



Se, por exemplo, se $\mathcal{H}_n = 71$, a medida de posição categórica é *bom*. O fato de uma pergunta ter 5 categorias e outra 4 não causa quaisquer problemas para a nota final, pois, quaisquer que sejam os índices calculados para as perguntas, a nota final será sempre a média aritmética dos índices.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BRASIL. Ministério da Saúde. Secretaria-Executiva. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. *Glossário temático: economia da saúde*. Brasília: Editora do Ministério da Saúde, 2005.
2. BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Gestão Pública. Programa Gespública. *Pesquisa de Satisfação: conceitos e orientações*. Brasília: Ministério do Planejamento, Orçamento e Gestão, 2013.
3. BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Gestão Pública. Programa Gespública. *Manual de Avaliação de Satisfação do Usuário do Serviço Público*. Disponível em <http://www.gespublica.gov.br/Tecnologias/anexos/instrumento_padrao_de_pesquisa_de_satisfacao/apostila_ips_jun10.pdf> Acesso em 20 jul. 2013.
4. ESPERIDIÃO, Monique Azevedo; TRAD, Leny Alves Bomfim. Avaliação de satisfação de usuários: considerações teórico-conceituais. *Cad. Saúde Pública*, v. 22, n. 6, p. 1267-1276, 2006.
5. Edwards, A. (1957): *Techniques of Attitude Scale Construction*. Appleton-Century-Crofts, New York.
6. Institute of Medicine (1990): *Medicare: a Strategy for Quality Assurance*, vol. 2. National Academy of Science.
7. Mclver, J. & E. Carmines (1981): *Unidimensional Scaling*. Sage Publications, Beverly Hills.
8. Ozcan, Y. (2009): *Quantitative Methods in Health Care Management: Techniques and Applications*. 2nd edition, Jossey-Bass, San Francisco.
9. Silva, N. (1988): *Amostragem Probabilística*. EDUSP, São Paulo.

10. Souza, J. (1988): *Métodos de Escalagem Psicossocial (Uni e Multidimensional)*, Métodos Estatísticos nas Ciências Psicossociais, vol. V. Thesaurus, Brasília.