



UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS  
DEPARTAMENTO DE SOCIOLOGIA

*4.3 (VI)*  
*Conservação*

# TABULAÇÃO CRUZADA EM PESQUISA SOCIAL

**Reginaldo Prandi**

(Professor Titular de Sociologia  
da Universidade de São Paulo)

Para uso didático exclusivo  
da disciplina FSL203/FFLCH/USP  
Versão 3

São Paulo, 1995

## Sumário

<b>Capítulo 1: Introdução à tabulação cruzada.....</b>	<b>2</b>
1.1. Pesquisa qualitativa e pesquisa quantitativa.....	2
1.2. Hipóteses como relações simétricas ou assimétricas.....	3
1.3. Definição nominal e definição operacional.....	5
1.4. Variáveis qualitativas e quantitativas.....	8
1.5. Modelos simples para hipóteses com duas variáveis.....	10
<b>Capítulo 2: Porcentagens em tabelas de duas entradas.....</b>	<b>13</b>
2.1. Tabulando duas variáveis, comparando proporções.....	13
2.2. Tabelas mudas.....	14
2.3. Computando as frequências absolutas das caselas.....	17
2.4. Como calcular as porcentagens.....	20
2.5. Apresentação das tabelas de duas entradas.....	32

## **Capítulo 1:**

### **Introdução à tabulação cruzada**

#### **1.1. Pesquisa qualitativa e pesquisa quantitativa**

Na pesquisa social, quando se estuda um grupo pequeno, uma população com a qual se possa conviver, por exemplo, interagindo com os seus membros, conhecendo a todos, conversando, entrevistando, observando, o próprio dia-a-dia da pesquisa (sistematizado pelo pesquisador-observador segundo roteiros previamente traçados) pode se constituir numa fonte quase que direta de informações buscadas pela pesquisa.

O pesquisador registra essas informações em seus "diários de campo", fitas magnéticas, fotografias, gerando um acervo qualitativo que será sua principal fonte de referência empírica nos momentos de análise daquelas situações pesquisadas. Esses procedimentos de campo são usuais sobretudo na antropologia, embora possam ser aplicados também em outras ciências sociais. Assim pode proceder um sociólogo ao estudar, por exemplo, uma rua, uma favela, um cortiço, um bar, um clube, uma escola, um hospital. O pesquisador aí quase sempre se sentirá como se estivesse na "sua tribo", pois a proximidade entre o pesquisador e o objeto é bastante grande.

Por outro lado, há muitas situações em que se trabalha com populações ou universos numericamente grandes. Essas populações grandes podem ser constituídas de indivíduos, instituições, aglomerados populacionais, jornais, documentos etc. Em todos estes casos, dificilmente o pesquisador tem a possibilidade de estudar diretamente cada uma das unidades que compõem a população ou é capaz de reter as informações sem o uso de algum meio especialmente desenvolvido para armazenar e sistematizar os dados colhidos. Quando se investiga a população de um bairro grande, de uma cidade, um colégio eleitoral, a rede de escolas de um estado, uma coleção bastante abrangente de

números de um dado jornal, um arquivo repleto de processos judiciais, e assim por diante, a forma mais fácil (talvez a única viável) consiste em transformar as informações de modo a poder tratá-las numericamente.

Por exemplo, pode-se contar quantos indivíduos podem ser incluídos numa categoria de um esquema de classificação, quantos se enquadram numa segunda categoria, quantos numa terceira etc. O pesquisador não terá que fazer esta contagem pessoalmente, podendo trabalhar com uma equipe preparada para isto. O uso de questionários e formulários padronizados permite esse tipo de coleta. E este material, depois, poderá ser submetido a modelos de cálculos de que o pesquisador pode dispor para sintetizar o conjunto de toda a informação colhida caso a caso.

Este segundo conjunto de procedimentos, que envolve necessariamente algumas noções básicas de estatística, é em geral chamado de pesquisa quantitativa.

No momento presente, o que nos interessa são exatamente estas situações de pesquisa em que estamos trabalhando com populações grandes, abordadas através de instrumentos de coleta padronizados do tipo do questionário, tendo como objetivo a verificação de hipóteses relacionais.

## **1.2. Hipóteses como relações simétricas ou assimétricas**

Hipótese relacional é uma proposição que estabelece, dentro de um campo teórico específico, a existência de alguma capacidade preditiva de uma dimensão ou aspecto do objeto de estudo com relação a outra dimensão ou aspecto. (Poderíamos estar tratando de vários aspectos ao mesmo tempo, mas para efeito didático, vamos nos manter por ora com apenas dois.) Pode ser que dois aspectos variem um com relação ao outro sem que a teoria nos informe sobre a existência de uma ordem de precedência de um sobre o outro. Numa outra situação, é fácil imaginar que muitos conteúdos teóricos são capazes de organizar hierarquicamente os fatos e acontecimentos, de tal sorte que um pode ser explicado, pelo menos parcialmente, pela variação do outro.

No primeiro caso, trata-se de um modelo simétrico - não importa qual influencia qual, mas sim que ambos os aspectos considerados seguem conjuntamente um padrão de variação. No segundo caso, em que um dos característicos é visto como resultante ou efeito de um outro, trata-se de um modelo assimétrico.

Pensemos em exemplos simples. Pessoas que ainda acham que a mulher não deve trabalhar fora de casa tendem também a defender a prática sexual antes do casamento apenas para os homens. Ambas as atitudes fazem parte de um mesmo quadro de tradicionalismo, o qual se compõe de atitudes, valores e modos de agir que abrangem desde o mais íntimo aspecto da vida privada até as expressões da vida pública de maior visibilidade. Em geral, pessoas de mentalidade tradicional tendem também a ser autoritárias, isto é, não democráticas, mostram-se preconceituosas com relação a diferenças raciais, religiosas etc. Tudo isto compõe aquilo que se chamou de "personalidade autoritária".<sup>1</sup> Neste quadro teórico, podemos dizer que o que se pensa do trabalho da mulher e do sexo pré-conjugal compõem uma relação simétrica, ou seja, considerados estes dois aspectos, não faz sentido dizer que um tem precedência sobre o outro, que um "causa" o outro. A sociologia tem demonstrado, porém, que indivíduos provenientes de meios sociais mais em contato com o processo de mudança social, mais informados, mais escolarizados, são menos tradicionais a respeito dos papéis femininos e do comportamento sexual. Em outras termos, maior exposição às instituições educacionais corresponde a uma patamar mais elevado de tolerância em relação ao outro.

Neste caso, se relacionarmos escolaridade com tolerância para com certas atitudes e práticas sociais, o critério escolaridade pode ser entendido como fonte de variação das atitudes e dos modos de agir.<sup>2</sup> Seguindo o exemplo, a idealização do lugar da mulher da sociedade é um elemento *dependente* do fator educacional que, por contraposição, é designado *independente*.

---

<sup>1</sup> Ver T.W. ADORNO *et alii*, *The Authoritarian Personality*. New York, Harper & Brothers, 1950.

<sup>2</sup> O nível de escolaridade indica, além do patamar de educação formal alcançado pelo indivíduo, que ele teve contato sistemático com um grupo específico em que as idéias valorizadas não são necessariamente as do grupo familiar.

### 1.3. Definição nominal e definição operacional

Lembremo-nos que estamos diante da situação em que nos vemos preocupados em estabelecer um meio para verificar se aceitamos, ainda que temporariamente, uma hipótese como verdadeira, ou se a rejeitamos como falsa. O meio que buscamos aqui para se fazer esse teste é, de antemão, quantitativo.

Ora, as hipóteses estabelecem relações entre (pelo menos) dois aspectos, característicos ou dimensões do objeto. Estes são referidos através de construções conceituais. No âmbito da teoria, todo conceito científico, para ser reconhecido como tal, deve ser teoricamente explicitado, isto é, deve dispor de sua definição nominal, que é a maneira de dizer exatamente o que se entende por aquela expressão ou termo.

Nos estudos sobre personalidade autoritária, conduzidos no pós-guerra, nos Estados Unidos, por Adorno e colaboradores,<sup>3</sup> o autoritarismo é definido como uma predisposição que certas pessoas desenvolvem no sentido de defender-se e conformar-se acriticamente em relação a certas normas e fontes de autoridade que acreditam imutáveis e legítimas, de tal modo que, ao agirem assim, elas deixam de se sentir ameaçados pelas mudanças possíveis e pelas diferenças. Assim, essas pessoas submetem sua capacidade individual de avaliação e adaptação a novas situações e contextos sócio-culturais a modelos de pensamento e ação rigidamente estereotipados, etnocêntricos e rejeitadores da expressão do outro.

Esse tipo de personalidade é frágil e por isto depende de uma autoridade exterior na qual crê, e na qual confia para afastar o temor e a ansiedade em relação ao perigo que lhe representa o outro desconhecido ou indesejado, podendo ser esta autoridade a própria tradição, ou então um líder, uma instituição, ou algum movimento ou partido que defenda tal tradição. Forma-se assim um sentimento de confiança numa ordem pré-estabelecida, sentimento evidentemente contrário a mudanças na sociedade. Esses sentimentos desenvolvem-se e se cristalizam no interior de grupos familiares e outras instituições sociais, o que pressupõe uma dependência da personalidade para com a estrutura social. A

---

<sup>3</sup> *Ibidem*, Chapter 1 *et passim*.

própria idéia de mudança e sua percepção dependem da estrutura da sociedade. Indivíduos de personalidade autoritária tendem a reunir as coisas em categorias bem rígidas, e este é o princípio da conformação do preconceito, pois às vezes basta um indício para que toda uma heterogeneidade se transmute, na mente do preconceituoso, numa só coisa, em geral negativa, desprezada e temida - daí poder-se seguir a discriminação, que já não reflete tão somente um sentimento, mas constitui-se de ações concretas.

Para efeito de pesquisa empírica, todo conceito, partindo da definição nominal, deve ser traduzido numa definição operacional, ou seja, numa proposição de indicadores selecionados para permitir a classificação ou mensuração da característica de que trata o conceito em cada uma das unidades a serem pesquisadas. Indicadores são referências, pistas, dicas, indícios sobre o aspecto considerado.

Dependendo do quadro teórico, a renda familiar pode ser um indicador de classe social; numa outra teoria, talvez se deva lançar mão da relação de produção para identificar a classe; numa outra perspectiva, ainda, talvez seja importante considerar também o *status* ocupacional do indivíduo. Quero dizer que não é possível encontrar indicadores se não se dispõe de um conceito e que estes não fazem sentido a não ser quando referidos a uma dada teoria.

A definição operacional inclui também os procedimentos que reúnem diferentes indicadores na composição de uma única medida ou padrão de medida para aquele aspecto da realidade que nos interessa. Podemos usar vários itens que compõem as comodidades de um domicílio para classificar um indivíduo ou família numa escala de estratificação social, e tanto a seleção desses itens como o peso que cada um deve receber fazem parte da definição operacional. Uma vez que essas possibilidades de escolha mudam com o progresso tecnológico, hábitos de consumo, padrões de mercado etc., não é difícil aceitar que estes passos todos estão carregados de historicidade, o que nos obriga, nas ciências sociais, a refazer sempre as definições operacionais, não se podendo trabalhar com um padrão único universal, mesmo quando lidamos com realidades aparentemente muito próximas.

Quando estamos estudando o que muitos chamam de "estados internos", como valores, expectativas, motivações, opiniões etc., é necessário construir indicadores que sejam capazes de traduzir essa "interioridade" em algo que possa ser "objetivado", "exteriorizado", "visível". Para isto os questionário com suas baterias de perguntas são bastante importantes. Parte-se do princípio de que os "estados internos" podem se manifestar ao pesquisador na forma de uma reação verbalizada a um estímulo também verbal (e às vezes visual): respostas dadas pelo pesquisado a perguntas formuladas pelo pesquisador.

São muitas as estratégias desenvolvidas na construção dos questionários e não vou tratar disto agora. Mas gostaria de dar uns exemplos. Numa pesquisa sobre anti-semitismo na Argentina<sup>4</sup>, foram aplicadas quatro perguntas: 1) "Em sua opinião, quais são as pessoas ou grupos mais prejudiciais a este país?" Esta pergunta era "aberta", ou seja, não se oferecia nenhuma alternativa prévia ao pesquisado. 2) "Quais as pessoas ou grupos que trazem mais benefícios para o país?" era uma segunda pergunta. 3) Em seguida se apresentava uma lista de doze grupos, entre os quais "judeus", "políticos", "portenhos", "grevistas", e se perguntava quais deles eram prejudiciais ou benéficos para o país. 4) Depois apresentava-se uma lista com dez "nacionalidades" (ingleses, franceses, judeus etc), perguntando-se quais deveriam ser atraídos ou impedidos como imigrantes para o país. Deste modo, o questionário criava condições de captar o preconceito anti-semita mais arraigado e explícito, que é aquele que se mostra mesmo quando o estímulo é mais fraco por não mencionar o grupo em questão, até aquelas situações de preconceito mais tênue, exigindo então um estímulo mais direto.

A construção de escalas (que permitem obter um gradiente, uma ordenação da intensidade) também tem sido muito usada e um exemplo fácil é a chamada escala de distância social, elaborada originalmente nos anos 20 por Bogardus para estudar relações raciais nos Estados Unidos. Versões desta escala têm sido usadas até hoje para medir

---

<sup>4</sup> Cf. Gino GERMANI, *Antisemitismo ideológico y antisemitismo tradicional. Comentario*, Buenos Aires, IX (34), 1962. Trata-se de *survey* realizado em Buenos Aires com uma amostra de 2078 informantes.



preconceito em relação a grupos e minorias. Vejamos um exemplo aplicável ao preconceito contra homossexuais. Você pode perguntar, para cada uma das situações seguinte, se o entrevistado se sentiria "muito constrangido", um "pouco constrangido", "absolutamente nada constrangido", ou "indiferente":

1. sabendo que homossexuais moram no seu bairro
2. tendo um homossexual como colega de trabalho
3. tendo um homossexual como companheiro de estudo
4. tendo um homossexual na família
5. tendo um homossexual como companheiro de lazer

Claro que esta escala de distância social depende das formas de preconceito e discriminação em relação ao grupo existentes no meio social pesquisado, sendo por vezes difícil estabelecer o gradiente, isto é, chegar à ordenação das situações como indicador de intensidade. Neste exemplo, a resposta "muito constrangido" na situação 1 revela o nível mais elevado de preconceito (morar no mesmo bairro não implica necessariamente interagir com eles); "nada constrangido" na situação 5 (que envolve a idéia de escolha, de livre opção, numa atividade que pressupõe intimidade - o oposto da situação 1 daria o nível mais baixo do preconceito, ou sua ausência, nos limites deste indicador.

Quando os esquemas de definição operacional estão concluídos, o seu resultado pode ser expresso ou em números que indicam quantidades, ou em categorias que indicam pertencimento a grupos assemelhados. E assim chegamos, finalmente àquilo que se chama variável.

#### **1.4. Variáveis qualitativas e quantitativas**

Variável é uma característica de um objeto, fenômeno, ou fato, que pode se apresentar diferente em cada caso observado, ou para grupos de casos, e que já passou pela etapa de operacionalização, podendo assim fornecer resultados que possam alimentar modelos numéricos de sistematização e síntese. A variável é uma reconstrução de algo do mundo, imediato ou não, a partir de um sistema de convenção capaz de dotar aquilo que é classificado ou medido de um critério de objetividade aceito pelo conjunto de pessoas interessadas naquele problema.

Uma variável é dita variável quantitativa quando ela oferece resultados numéricos que podem ser somados, subtraídos, divididos, multiplicados.<sup>5</sup> As variáveis quantitativas são raras nas ciências sociais, e em geral são tomadas de empréstimo da economia, da demografia, da psicologia, como, o tamanho das famílias de determinada população, a renda monetária das unidades domésticas, a produtividade física de estabelecimentos agrícolas, o número de anos de escolaridade formal dos indivíduos, as taxas de evasão escolar em bairros de uma cidade, a nota atribuída num teste psicológico. Elas são resultados: 1) de contagens (ou de operações feitas entre cálculos de aspectos distintos, por exemplo a renda per capita, em que um montante de dinheiro é dividido por um número de pessoas), ou 2) da comparação com um padrão de mensuração capaz de fornecer uma quantia mínima como uma unidade de medida (caso do comprimento, área, volume, massa, peso, tempo, temperatura...)<sup>6</sup>.

Uma variável é dita variável qualitativa ou nominal quando ela fornece categorias de classificação que não podem ser submetidas às operações aritméticas; os resultados não exprimem quantidades, magnitudes, mas sim expressam diferenças e semelhanças, permitindo separar (para comparar) as unidades de pesquisa em diferentes grupos, internamente homogêneos, a que chamamos de atributos, características, alternativas, classes, categorias. A única operação numérica aplicável às variáveis qualitativas

---

<sup>5</sup> Uma variável quantitativa pode ter um zero absoluto ou arbitrário. Fala-se em zero absoluto quando, ao assumir este valor, não há mais nenhuma quantidade do fenômeno que está sendo medido. Estas variáveis são chamadas escalas de razão e representam o nível de mensuração mais elevado. Como exemplo temos as variáveis métricas. Quando o zero é arbitrário, o fenômeno não se anula no ponto zero, como nas escalas Celsius e Fahrenheit de temperatura. Quando conseguimos obter variáveis quantitativas para medir estados internos (atitudes, valores), elas são sempre deste segundo tipo, que recebe o nome de escalas de intervalo. De outro ponto de vista, as variáveis são chamadas discretas, quando só podem assumir valores inteiros (como número de filhos); ou variáveis contínuas, quando frações da unidade fazem parte dos valores previstos.

<sup>6</sup> Note que quando a variável é expressa em unidades de medida, esta unidade faz parte do número e assim, quando se executa uma operação aritmética com o número, a unidade também sofre o efeito do cálculo. Um comprimento em centímetro multiplicado por si mesmo, gera um número que não expressa mais um comprimento, mas uma área. Muda-se de um conceito para outro quando se faz, por exemplo, a seguinte conta:  $20 \text{ cm} \times 20 \text{ cm} = 400 \text{ cm}^2$ . Este simples exemplo mostra como a matemática é importante no corpo teórico das ciências mais formalizadas. Uma segunda observação diz respeito ao fato de que medidas padronizadas estão organizadas em sistemas resultantes de convenções. Um problema sério para as ciências, mesmo as mais formalizadas, como a física teórica, é que elas não dispõem de um sistema único

consiste em classificar as unidades pesquisadas e contar quantas delas pertencem a cada uma das diferentes categorias. Assim, o número que é estatisticamente manipulado é a frequência numa dada categoria da variável ou a frequência numa combinação de categorias de mais de uma variável.<sup>7</sup>

### 1.5. Modelos simples para hipóteses com duas variáveis

Se levamos em conta que uma hipótese relacional entre dois característicos de um objeto pode ser:

- a) ou do tipo de relação simétrica ou assimétrica,
- b) com variáveis ou qualitativas ou quantitativas,

podemos antever a possibilidade de várias combinações que podem surgir numa situação de pesquisa. Vou tentar resumir isto no Quadro 1.

No Quadro 1 estão dados oito diferentes modelos simples, embora haja alguns outros possíveis<sup>8</sup>. De todos eles, o que mais se usa em ciências sociais é o modelo 1, a tabulação cruzada para variáveis qualitativas, pressupondo-se que uma pode afetar a outra.

Esta preponderância de um modelo tão simples como o modelo 1 decorre do fato de que, em geral, nós trabalhamos basicamente nas ciências sociais com nível de mensuração nominal (nossas variáveis são "fracas" do ponto de vista do rigor da mensuração) e de que nossos modelos explicativos buscam estabelecer relações de causalidade ou influência entre esferas ou dimensões diferentes da realidade social.

---

capaz de dar conta de todos os aspectos de seus objeto. No âmbito teórico que precede essas etapas operacionais, o mesmo pode ser dito em relação às ciências sociais.

<sup>7</sup> Quando a variável qualitativa tem apenas duas categorias (como homem ou mulher, jovem ou velho, economicamente ativo ou inativo, nativo ou migrante, analfabeto ou alfabetizado, criado no campo ou na cidade, rico ou pobre, autoritário ou tolerante, líder ou liderado, e assim por diante), podemos atribuir a uma das categorias o valor 1 e à outra o valor zero, e tratar esta variável como se ela fosse quantitativa. Este tipo de variável é chamado de variável *dummy*.

<sup>8</sup> Por exemplo, se as variáveis são qualitativas, mas permitem algum ordenamento (como estratificação social, escalas de prestígio, graus de escolaridade formal), há modelos especiais de regressão e de correlação de postos. Entretanto, ninguém errará se usar um dos modelos propostos, mesmo nestes casos, embora possa perder informação do ponto de vista estatístico.

**Quadro 1**

**CRUZAMENTO DE DUAS VARÁVEIS  
ESQUEMA DOS MODELOS SIMPLES MAIS USADOS**

<b>a) Relações assimétricas (a VI afeta a VD)</b>		
Váriável Independente VI	Váriável dependente VD	
	Qualitativa	Quantitativa
Qualitativa	modelo 1. Tabulação cruzada: comparação de porcentagens calculadas sobre o total de cada categoria da VI	modelo 2. Comparação de médias da VD calculadas para cada categoria da VI
Quantitativa	modelo 3. Agregam-se valores da VI em faixas e aplica-se o modelo 1	modelo 4. equações de regressão: simples para uma VI múltiplas para mais de uma VI
<b>b) Relações simétricas (não há precedência entre as variáveis Vx e Vy)</b>		
Váriável Y Vy	Váriável X Vx	
	Qualitativa	Quantitativa
Qualitativa	modelo 5. Tabulação cruzada: Comparação de porcentagens calculadas sobre o total geral da tabela; coeficientes de associação	modelo 6. Agregar valores de Vx em faixas e aplicar o modelo 5
Quantitativa	modelo 7. Agregar valores da Vy em faixas e aplicar o modelo 5	modelo 8. Coeficientes de correlação

No Capítulo 2 veremos especificamente a situação em que temos duas variáveis qualitativas (ou transformadas em qualitativas), das quais, de acordo com a hipótese que orienta a análise dos dados empíricos, uma delas é tomada como a variável que sofre efeito (variável dependente) produzido pela outra (variável independente). Talvez seja importante chamar a atenção para o fato de que esta idéia de "causa e efeito" pode ter sentido meramente formal, uma vez que a causalidade pode estar fora desta situação simples de dois aspectos, numa trama de causas, influências, relações bastante mais abrangente.

Em Durkheim, para tomarmos o exemplo fundante de *O suicídio*, o estado conjugal, a ocupação, a religião, a idade, as estações do ano, o país etc. são tomados como variáveis independentes que indicam a possibilidade maior ou menor de ocorrência do suicídio, que é a variável dependente, a variável objeto que se pretende explicar. Mas na teoria durkheimiana, a causa buscada é de ordem sociológica: a fonte de causalidade está nas forças sociais que controlam a vida dos indivíduos. Essas variáveis independentes são meros indicadores para graus mais fortes ou mais fracos de interação social e coersão, assim como o ato do suicídio indica, nas sociedades de solidariedade orgânica, uma ruptura extrema do indivíduo para com a ordem social. As instituições sociais controlam também o imperativo de viver. Deste modo, aqueles indicadores refletem a força de instituições e papéis sociais que atuam no controle do comportamento, organizando forças sociais que estão acima dos indivíduos, forças, por conseguinte, sociológicas.

Poder-se-ia dizer, portanto, que o trabalho do cientista social só tem sentido quando a relação causa-efeito que ele estuda nesta modalidade de pesquisa está explicada num todo teórico capaz de organizar todas essas relações parciais, lembrando-nos que um procedimento de pesquisa empírica não tem valor em si mesmo, uma vez que a pesquisa é apenas um meio para a construção da explicação, construção da teoria.

Também, quando se fala em causa-feito, pode-se estar simplesmente fazendo referência a uma situação em que um ou uma série de fatores contribui, com direção e

intensidade diversas, para que determinado evento assuma este ou aquele resultado, forma ou conteúdo.

## **Capítulo 2:**

### **Porcentagens em tabelas de duas entradas**

#### **2.1. Tabulando duas variáveis, comparando proporções**

Vamos a partir de agora seguir os passos necessários à verificação de hipóteses como nos propusemos. Usarei tabelas especialmente preparadas para este fim, com dados verídicos, fornecidos por um *survey* sobre a questão da Democracia Delegativa, realizado em São Paulo, com aplicação, em dezembro de 1991, de um questionário a uma amostra de 800 eleitores que residem e votam na Região Metropolitana da Grande São Paulo. A pesquisa é uma realização do CEBRAP (Centro Brasileiro de Análise e Planejamento), sob direção de Guillermo O'Donnell. A pesquisa tem por fim testar hipóteses de O'Donnell sobre o tipo de democracia idealizada pelos eleitores paulistas.<sup>9</sup> Como é costume, e por razões práticas, o *survey* investiga várias dimensões da vida pública e privada dos informantes, coletando dados para além daqueles necessários à verificação das hipóteses de O'Donnell.

Para o presente capítulo selecionei algumas variáveis sem nenhuma preocupação com coerência teórica, uma vez que aqui estou apenas interessado em mostrar como proceder no uso de tabelas de duas entradas.

---

<sup>9</sup> Ver Guillermo O'DONNELL, *Democracia Delegativa? Novos Estudos Cebrap*, São Paulo, no. 31, pp. 25-40, outubro 1991.

## 2.2. Tabelas mudas

A partir de uma hipótese, devemos preparar o material para sua verificação. Suponhamos que estou interessado em saber se indivíduos criados em meios sociais mais tradicionais, relativamente distantes dos efeitos modernizantes dos processos de mudança social que se dão no desenvolvimento da sociedade capitalista contemporânea, se estes indivíduos, por assim dizer "mais isolados", tendem a uma mais fraca internalização de valores que emergem do processo de constituição da cidadania que vem junto com o desenvolvimento da sociedade de classes moderna. Vou tomar como indicador de maior grau de isolamento na origem social o fato de o indivíduo ter sido criado até o 16 anos de idade em zona rural (em oposição a cidade), e como indicador de fraca internalização de princípios de direitos da pessoa humana a opinião concordante com a legitimidade de se torturarem criminosos. Tal concordância nos indicará a presença de mentalidade tradicionalista, intransigente, anti-democrática ou autoritária.

Tenho uma hipótese, tenho os indicadores. Minha formulação afirma que eu suponho que o meio social de origem afeta a formação da opinião. Tenho portanto uma variável dependente - a opinião - e uma variável independente - a origem. Ambas são variáveis qualitativas, pois sobre a origem apenas tenho a informação do local onde foi criado, e para a opinião disponho tão somente de respostas verbais do entrevistado a uma pergunta que lhe foi apresentada pelo entrevistador. E já sei que o modelo a ser usado é o de comparação de porcentagens para as classes da VI, que no Capítulo 1 chamei de modelo 1. Falta completar a tarefa.

Quando nos propomos a trabalhar um cruzamento, já sabemos quais são as variáveis envolvidas e quantas categorias tem cada uma delas. Assim, o número de combinações (categorias cruzadas) possíveis é igual à multiplicação do número de classes da VI (variável independente) pelo da VD (variável dependente). Podemos então desenhar o esqueleto da tabela, ainda sem as frequências absolutas das caselas, ou seja o número de casos classificados em cada combinação entre uma categoria de uma variável e uma categoria da outra. Estes "esqueletos" são chamados "tabelas mudas": elas não dizem nada, mas podem orientar a tabulação.

Uma tabela de duas entradas é formada por linhas e colunas. Se uma variável está disposta de tal modo que cada uma de suas categorias encabeça uma coluna da tabela, ela é dita variável das colunas. Se cada categoria inicializa uma linha da tabela, a variável é chamada variável das linhas.

Examine a Tabela 1 apresentada a seguir. Nas colunas está a variável sobre o local de residência do entrevistado até os 16 anos de idade (P68)<sup>10</sup>, sendo que "cidade" encabeça a primeira coluna e "campo", a segunda. Nas linhas aparecem as quatro categorias da variável de opinião sobre torturar criminosos (P62D). A Tabela 1 tem, assim, oito caselas, as quais deverão ser preenchidas com a tabulação conjunta das duas variáveis.

Mas podemos dispor as mesmas informações invertendo a posição das variáveis, como mostra a Tabela 2. Nesta, a origem foi para as linhas e a opinião para as colunas. Nada mudou substantivamente. Trata-se apenas de uma decisão estética, ou prática, se você imagina como a tabela ficará numa página do seu trabalho. Desenhar as tabelas mudas ajuda a decidir qual das duas formas fica melhor, mais fácil de visualizar, mais apropriado para digitar etc. Questão de gosto. Algumas editoras e revistas científicas podem impor ao pesquisador o gosto das suas comissões editoriais. Não faz diferença no conteúdo.

Mesmo depois de feita a tabulação dos dados, a tabela pode ser reorganizada de modo a se chegar a um melhor desenho. O que importa é que os números apresentados sigam as regras formais de verificação de hipótese.

---

<sup>10</sup> Esta sigla da variável segue exatamente a enumeração usada no questionário original e ela foi usada como referência pelo sistema de tabulação. Assim, após definir as variáveis, quando me interessava obter o cruzamento mostrado na Tabela 1, bastava usar no sistema SPSS/PC o comando "Crosstabs P62D By P68". Mantenho a sigla aqui apenas para que o leitor iniciante nesse tipo de procedimento vá se familiarizando com as mensagens impressas.



Tabela 1 (\*). TABELA MUDA (tabela sem as frequências das caselas)

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
By P68 MOROU NA ROCA OU CIDADE ATE OS 16

P68->		Cidade 1	Zona ru- ral 2	Total
P62D		coluna 1	coluna 2	Total
(linha 1)	1			
Concorda Total				
(linha 2)	2			
Depende do Caso				
(linha 3)	3			
Discorda Total				
(linha 4)	9			
Nao Sabe				
Total				

Observação: Os números que aparecem junto ao nome de cada categoria são os códigos usados no questionário e que foram digitados para o armazenamento dos dados num arquivo para facilitar a tabulação por computador, ou manual até. Estes números não são quantidades: representam categorias qualitativas!

Tabela 2 (Tabela muda)

Crosstabulation: P68 MOROU NA ROCA OU CIDADE ATE OS 16  
By P62D CONCORDA TORTURA DE CRIMINOSOS

P62D->		Concorda Total 1	Depende do Caso 2	Discorda Total 3	Nao Sabe 9	Total
P68						
Cidade	1					
Zona rural	2					
Total						

(\* Os dados aqui apresentados são provenientes da Pesquisa sobre Democracia Delegativa, realizada pelo CEBRAP, sob direção de Guillermo O'Donnell. Trata-se de um survey para a Região Metropolitana da Grande São Paulo, aplicado em dezembro de 1991.

### 2.3. Computando as frequências absolutas das caselas<sup>11</sup>

Decidido qual formato de tabela usar, temos que proceder à tabulação cruzada propriamente dita, isto é, computar as frequências conjuntas segundo os dois critérios de classificação propostos.

Isto, se fosse feito manualmente, consistiria, por exemplo, em separar os questionários em duas pilhas: os da cidade e os da roça. Depois, para cada pilha separaríamos os que são totalmente de acordo com a tortura de criminosos, os que concordam dependendo do caso, os que discordam totalmente e os que não sabem (não têm opinião formada). Teríamos assim oito pilhas. Basta contar quantos questionários há em cada uma e teremos a frequência absoluta de cada casela. Hoje há vários sistemas bastante práticos para tabulação eletrônica.

Se você escolheu a Tabela 1, os resultados da tabulação cruzada estão na Tabela 3. Se preferiu a Tabela 2, consulte as frequências cruzadas na Tabela 4. Compare uma com a outra e veja como, de fato, os resultados são os mesmos:

160 criaram-se na cidade e concordam com a tortura,

150 foram criados na cidade e concordam com a tortura em certas situações,

42 cresceram em zona rural e concordam em certos casos,

60 foram socializados na roça e discordam totalmente,

... e assim por diante, em ambas as tabelas.

As somas das frequências das caselas nas direções das linhas e das colunas reproduzem as frequências de cada uma das variáveis tomadas separadamente, e que são obtidas também através da tabulação simples que se faz para cada uma das variáveis do estudo, muito antes de se começar a trabalhar com tabulações cruzadas.

---

<sup>11</sup> As tabelas deste capítulo são provenientes da Pesquisa sobre Democracia Delegativa, realizada pelo CEBRAP, sob direção de Guillermo O'Donnell. Trata-se de um *survey* para a Região Metropolitana da Grande São Paulo, aplicado em dezembro de 1991.

Segundo a Tabela 3 ou a Tabela 4, uma das somas das caselas diz que temos 626 eleitores criados na cidade e 168 no campo, respectivamente, 78,8% e 21,2% da amostra válida para esta tabela. Amostra válida é a amostra total excluídos, porém, os casos para os quais faltou algum tipo de informação necessária para a classificação nesta tabela cruzada; de fato, há aqui seis casos de informação faltando, *missing observations*.<sup>12</sup>

Na outra margem das referidas tabelas, temos 223 casos (ou 28,1%) com total concordância com a tortura, 192 (24,2%) com posição a depender da situação, 361 (45,5%) totalmente discordantes, e 18 pesquisados (2,3%) que não souberam (ou não quiseram) opinar.<sup>13</sup> Estas duas distribuições, uma para cada variável, são chamadas distribuições marginais, pelo simples fato de virem dispostas nas margens da tabela.<sup>14</sup>

---

<sup>12</sup> Conhecer as distribuições simples de cada variável antes de organizar a tabulação ajuda a evitar que categorias da VI tenham menos que 25 casos, pois se isto ocorre, a análise estatística ficará prejudicada. Ao se trabalhar com uma VI que originalmente tem uma ou mais categorias com menos de 25 casos, há dois caminhos: 1) agregar categorias, se isto fizer sentido teoricamente, trabalhando-se então com números maiores; 2) não considerar as categorias da VI com menos de 25 casos para efeito de análise, mantendo-as na tabela numa forma que veremos mais adiante, ou atribuindo a elas a condição *missing observations*, que exclui os casos que nela se enquadrarem. Em muitas situações de pesquisa, a condição *missing data* leva a exclusão de casos da tabela, especialmente quando a questão "não se aplica" ao caso. Por exemplo, se você estiver tabulando as horas de trabalho para uma determinada população, entrarão em *missing observations* todos os casos em que se registrou que a pessoa não trabalha, logo, não há para ela horas a computar, e a questão não se aplica neste caso específico. Quando se trata de falta de resposta, cabe ao pesquisador incluir estes casos numa categoria "não respondeu" ou jogá-los em *missing observations*.

<sup>13</sup> Estes números podem diferir um pouco das distribuições obtidas originalmente para cada variável, visto que 6 casos foram excluídos da tabela por falta de informação (*missing data*) em pelo menos uma das variáveis.

<sup>14</sup> Segundo a teoria da probabilidade, as distribuições marginais fornecem as probabilidades de ocorrência dos eventos "cidade" ou "zona rural" para a variável local de criação, e as probabilidades de ocorrência das diferentes respostas consideradas sobre a questão da tortura. Se um indivíduo desta população é sorteado ao acaso, podemos dizer que há uma chance de 21,2% ou 0,212 dele ter sido criado na zona rural e uma chance de 45,5% ou 0,455 dele ser contrário à tortura de prisioneiro. Mas não sabemos ainda, antes de tabular os dados conjuntamente, qual seria a chance dele ter se criado na roça e, ao mesmo tempo, discordar da tortura. Ensina a teoria da probabilidade que, se não houver nenhuma influência entre estas duas variáveis, a probabilidade conjunta destes dois eventos será igual ao produto das probabilidades de cada um tomadas separadamente. Assim, neste exemplo, a probabilidade de encontrarmos alguém criado na roça e que seja ao mesmo tempo contrário à tortura seria:  $0,212 \times 0,455 = 0,096$  ou 9,6%. Como temos 794 casos válidos, 9,6% representam 77 casos. Se a frequência computada para esta casela for razoavelmente diferente de 77, diremos que as variáveis estão inter-relacionadas. Mas isto é uma outra história...

Tabela 3

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
By P68 MOROU NA ROCA OU CIDADE ATE OS 16

P68->	Count	Cidade	Zona ru- ral	Row Total	
		1	2		
P62D					
Concorda	1	160	63	223	
Total				28.1	
Depende do Caso	2	150	42	192	
Total				24.2	
Discorda	3	301	60	361	
Total				45.5	
Nao Sabe	9	15	3	18	
Total				2.3	
Column Total		626	168	794	
		78.8	21.2	100.0	

Nota: Expressões em inglês geradas pelo sistema SPSS/PC:  
Count = frequências  
Crosstabulation = tabulação cruzada  
By = por  
Row Total = total de cada linha  
Column Total = total de cada coluna  
Missing Observations = casos não considerados para entrar na tabela

Number of Missing Observations = 6

Tabela 4

Crosstabulation: P68 MOROU NA ROCA OU CIDADE ATE OS 16  
By P62D CONCORDA TORTURA DE CRIMINOSOS

P62D->	Count	Concorda Total	Depende do Caso	Discorda Total	Nao Sabe	Row Total
		1	2	3	9	
P68						
Cidade	1	160	150	301	15	626
Total						78.8
Zona rural	2	63	42	60	3	168
Total						21.2
Column Total		223	192	361	18	794
		28.1	24.2	45.5	2.3	100.0

Number of Missing Observations = 6

## 2.4. Como calcular as porcentagens

A análise dos resultados consiste em comparar as frequências. Se tomamos a origem rural-urbana como VI, isto é, como aquela que pode ter alguma influência sobre a VD (opinião), devemos comparar as respostas entre cidade e zona rural, pois é como se estivéssemos comparando duas coisas diferentes, para saber se elas provocam resultados diferentes. Então, vemos, pela Tabela 3 ou pela Tabela 4, que a concordância total com a tortura de criminosos foi observada em 160 eleitores criados na cidade, contra 63 criados na zona rural. Seria um absurdo concluir contrariamente à hipótese, sem antes considerarmos que os tamanhos dos grupos são diferentes e por isto aquelas frequências devem ser relativizadas: concordam com a tortura 160 dos 628 criados na cidade, ou seja  $160/628 = 0,256$  ou  $0,256 \times 100 = 25,6\%$ ; e concordam com a tortura 63 dos 168 criados na zona rural, ou seja,  $63/168 = 0,375$ , que é o mesmo que  $0,375 \times 100 = 37,5\%$ .

Quando dividimos a frequência absoluta pelo respectivo total do seu grupo, encontramos a frequência relativa, que estatisticamente equivale a transformar ambos os totais no valor 1 (a soma das frequências relativas é sempre igual a 1). Como é pouco prático trabalhar com números menores que 1, multiplicamos as frequências relativas por 100 e obtemos as frequências percentuais ou porcentagens ou taxas. Com isto, agora, tanto o grupo dos criados na cidade como o dos criados em zona rural passou a ter o mesmo total, isto é, 100%, ficando assim comparáveis. Ora a concordância com a tortura de criminosos é de 25,6% para os criados na cidade, contra 37,5% para os da zona rural, o que reforça a hipótese: a socialização em meio urbano desfavorece atitudes autoritárias. Faça você mesmo os cálculos para a categoria "discorda totalmente".

Note que, se nossa base é a Tabela 3, as porcentagens são calculadas para os totais das colunas, pois nesta tabela a VI encabeça as colunas. Se usamos a Tabela 4, então as porcentagens são calculadas para os totais das linhas, pois aí a variável independente tem suas categorias nas linhas. Não importa se a VI está nas linha ou colunas, mas é decisivo você acompanhar os cálculos na direção dela. Esta é a regra de ouro das tabulações

cruzadas bidimensionais com uma das variáveis tomada como resultante e a outra como fonte de variação: compare os grupos que são pensados como capazes de introduzir diferenças na outra dimensão considerada, equalizando, para tanto, seus tamanhos em 100,0%.

Vamos olhar agora a Tabela 5, que traz as mesmas variáveis. A VI está nas colunas e a VD, nas linhas. Dentro de cada casela aparecem 4 números, cujo significado está indicado no canto superior esquerdo da tabela, ou seja,

Count	= frequência absoluta
Row Pct	= porcentagem nas linhas
Col Pct	= porcentagens nas colunas
Tot Pct	porcentagem sobre o total geral.

---

|

Para a nossa hipótese, somente a terceira linha de cada casela, que traz a porcentagem calculada para o total da coluna correspondente, nos interessa. O resto será desprezado e não será transcrito no relatório, artigo, tese etc.

Caso a VI estivesse nas linhas, nós teríamos que usar as porcentagens dadas na segunda linha desta tabela. E ainda, se nosso modelo fosse simétrico, usaríamos o total geral e, portanto, as porcentagens que nesta tabela estão transcritas na última linha de cada casela.

Tabela 5 (\*\*)

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
 By P68 MOROU NA ROCA OU CIDADE ATE OS 16

P68->	Count Row Pct Col Pct Tot Pct	Cidade	Zona ru- ral	Row Total	OBSERVAÇÕES:
		1	2		
P62D					Count = frequências
1	160	63	223		
Concorda Total	71.7	28.3	28.1		Row Pct = Porcentagens sobre totais das linhas
	25.6	37.5			
	20.2	7.9			
2	150	42	192		Col Pct = Porcentagens sobre os totais das colunas
Depende do Caso	78.1	21.9	24.2		
	24.0	25.0			
	18.9	5.3			
3	301	60	361		Tot Pct = Porcentagens sobre o total geral da tabela
Discorda Total	83.4	16.6	45.5		
	48.1	35.7			
	37.9	7.6			
9	15	3	18		
Nao Sabe	83.3	16.7	2.3		
	2.4	1.8			
	1.9	.4			
Column Total	626	168	794		
	78.8	21.2	100.0		

Number of Missing Observations = 6

(\*\*) O sistema SPSS, *Statistical Package for the Social Sciences*, para uso em microcomputadores da linha PC ou compatíveis, existente em versões para DOS e para Windows, tem opções de imprimir todos os quatro tipos de frequências: a absoluta e as relativas para os totais das linhas, das colunas, ou sobre o total geral da tabela, ou imprimir apenas aquelas que o pesquisador seleciona em função de seu problema.

Na Tabela 6 e na Tabela 7, apenas as porcentagens úteis para a verificação da hipótese estão impressas. Quando nós mesmos fazemos os cálculos, não perdemos tempo com cálculos inúteis. Quando usamos um programa de computador, também podemos pedir a impressão só das porcentagens que interessam. Se não se tem a hipótese bem definida, costuma-se pedir "tudo", mas isto é pura perda de tempo.

Tanto com base na Tabela 6 ou na Tabela 7, verificamos que posições tradicionalistas estão mais associadas à origem marcada pelo meio rural. Dos criados na cidade, 25,6% são favoráveis à tortura de criminosos, enquanto que essa taxa sobe para 37,5% quando a origem é rural. Na categoria intermediária, em que a posição é avaliada caso a caso, não há praticamente diferença. Quando se trata de total discordância, temos 48,1% para os socializados na cidade e 35,7% para os que se criaram na zona rural. As diferenças encontradas não são muito grandes, pois haverá um número razoável de outros fatores que devem aí também interferir, mas elas são suficientes para não rejeitarmos a hipótese por ora. Mesmo nesta época das comunicações eletrônicas e de intensa interação campo-cidade, o campo ainda carrega algo daquilo que Marx chamava no fim do século passado de "idiotia da vida rural".



Tabelas 6

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
 By P68 MOROU NA ROCA OU CIDADE ATE OS 16

P68->	Count Col Pct	Cidade 1	Zona ru- ral 2	Row Total
P62D				
	1	160	63	223
Concorda Total		25.6	37.5	28.1
	2	150	42	192
Depende do Caso		24.0	25.0	24.2
	3	301	60	361
Discorda Total		48.1	35.7	45.5
	9	15	3	18
Nao Sabe		2.4	1.8	2.3
Column Total		626 78.8	168 21.2	794 100.0

Number of Missing Observations = 6

-----  
 Tabela 7

Crosstabulation: P68 MOROU NA ROCA OU CIDADE ATE OS 16  
 By P62D CONCORDA TORTURA DE CRIMINOSOS

P62D->	Count Row Pct	Concorda Total 1	Depende do Caso 2	Discorda Total 3	Nao Sabe 9	Row Total
P68						
	1	160	150	301	15	626
Cidade		25.6	24.0	48.1	2.4	78.8
	2	63	42	60	3	168
Zona rural		37.5	25.0	35.7	1.8	21.2
Column Total		223 28.1	192 24.2	361 45.5	18 2.3	794 100.0

Number of Missing Observations = 6

Formulemos agora a hipótese de que atitudes sobre os direitos da pessoa humana dependem ou são afetadas pela posição do indivíduo na hierarquia dos estratos sociais, uma vez que indivíduos que fazem parte de famílias em melhores condições de vida tendem a ter maior acesso à informação, dependem menos de fontes tradicionais de visão de mundo e acompanham mais facilmente os processos de mudança social.

O indicador aqui usado para estrato social é a renda familiar mensal, agrupada em faixas, e o indicador de tradicionalismo-modernismo continua sendo a opinião sobre torturar criminosos.

Os dados estão na Tabela 8. Faça você mesmo a verificação, escolhendo os números corretos e redigindo uma pequena conclusão.

Agora compare os dados que você selecionou com aqueles oferecidos nas Tabelas 9 e 10. De fato, ao passarmos das rendas mais baixas para as mais altas, as taxas de concordância total com a tortura vão caindo: 35,8%, 32,5%, 26,2% até 11,2%. No grupo de renda mais baixa, a taxa de adesão à tortura é três vezes maior que a do grupo de melhor renda. A categoria intermediária flutua em torno dos 25%, mas quando tomamos a resposta contrária à tortura, as porcentagens sobem significativamente. Enquanto na amostra total (ver o total das linhas, *Row Total*) o máximo de discordância obtido foi de 45,7%, quando isolamos os informantes classificados na faixa de renda mais alta, este número sobe para 63,3%, que é o dobro do número encontrado para os de renda inferior a 3 salários mínimos (33,8%).

Assim, por ora, mantemos de pé a conclusão de que a posição ocupada pelo indivíduo numa estrutura social forjada pela coexistência de diferenças sociais (que são materiais e simbólicas) modela, orienta, conduz, modifica, afeta minimamente seu modo de sentir e pensar.

Tabela 8

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
 By P72 RENDA FAMILIAR EM SALARIOS MINIMOS

		Count					
P72->	Row Pct	1 a 3 SM	4 a 7 SM	8 a 20 SM	mais de 20 SM	Row Total	
	Col Pct						
	Tot Pct	1	2	3	4	Total	
-----+-----+-----+-----+-----+-----+-----							
P62D	1	53	78	74	11	216	
Concorda	Total	24.5	36.1	34.3	5.1	28.1	
		35.8	32.5	26.2	11.2		
		6.9	10.2	9.6	1.4		
-----+-----+-----+-----+-----+-----+-----							
	2	42	54	66	25	187	
Depende do Caso	Total	22.5	28.9	35.3	13.4	24.3	
		28.4	22.5	23.4	25.5		
		5.5	7.0	8.6	3.3		
-----+-----+-----+-----+-----+-----+-----							
	3	50	103	136	62	351	
Discorda	Total	14.2	29.3	38.7	17.7	45.7	
		33.8	42.9	48.2	63.3		
		6.5	13.4	17.7	8.1		
-----+-----+-----+-----+-----+-----+-----							
	9	3	5	6		14	
Nao Sabe	Total	21.4	35.7	42.9		1.8	
		2.0	2.1	2.1			
		.4	.7	.8			
-----+-----+-----+-----+-----+-----+-----							
	Column Total	148	240	282	98	768	
	Total	19.3	31.3	36.7	12.8	100.0	

Number of Missing Observations = 32

Observação: Casela vazia indica que não foi encontrado nenhum caso nesta combinação; a frequência é zero.

Tabela 9

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
By P72 RENDA FAMILIAR EM SALARIOS MINIMOS

P72->	Count Col Pct	1 a 3 SM 1	4 a 7 SM 2	8 a 20 SM 3	mais de 20 SM 4	Row Total
P62D						
Concorda	1	53 35.8	78 32.5	74 26.2	11 11.2	216 28.1
Depende do Caso	2	42 28.4	54 22.5	66 23.4	25 25.5	187 24.3
Discorda	3	50 33.8	103 42.9	136 48.2	62 63.3	351 45.7
Nao Sabe	9	3 2.0	5 2.1	6 2.1		14 1.8
Column Total		148 19.3	240 31.3	282 36.7	98 12.8	768 100.0

Number of Missing Observations = 32

Tabela 10

Crosstabulation: P72 RENDA FAMILIAR EM SALARIOS MINIMOS  
By P62D CONCORDA TORTURA DE CRIMINOSOS

P62D->	Count Row Pct	Concorda Total 1	Depende do Caso 2	Discorda Total 3	Nao Sabe 9	Row Total
P72						
1 a 3 SM	1	53 35.8	42 28.4	50 33.8	3 2.0	148 19.3
4 a 7 SM	2	78 32.5	54 22.5	103 42.9	5 2.1	240 31.3
8 a 20 SM	3	74 26.2	66 23.4	136 48.2	6 2.1	282 36.7
mais de 20 SM	4	11 11.2	25 25.5	62 63.3		98 12.8
Column Total		216 28.1	187 24.3	351 45.7	14 1.8	768 100.0

Number of Missing Observations = 32

A Tabela 11 foi computada para submetermos, agora, a variável dependente que estamos analisando em função de uma outra que não é do tipo estrutural, porém comportamental. A suposição é que indivíduos que se interessam ou não pela vida pública, e assim o demonstram acompanhando mais ou acompanhando menos o que acontece na política (este é o indicador), têm posições diferentes em relação aos direitos da pessoa humana. Os mais interessados tenderiam a desprezar práticas como a tortura.

Analise a Tabela 11. Os dados já "limpos" estão nas Tabelas 12 e 13. Note que, nestas, já não aparecem as freqüências absolutas das caselas, pois elas de nada servem para o nosso intento. Os números confirmam a hipótese. Agora, para variar, quero chamar a atenção também para a categoria "não sabe". O "não sabe/não responde" em perguntas de opinião é um bom indício de desinteresse, alienação, marginalização, falta de informação, e suas taxas costumam ser maiores para os grupos sociais estruturalmente mais distantes dos processos decisórios e da sua discussão. Por hipótese reiteradamente não rejeitada, esses grupos se caracterizam pelo sexo feminino, pelos indivíduos idosos, pelos aposentados, os pouco escolarizados, aqueles provenientes de meios mais distantes dos grandes centros onde se forma a opinião pública do país, os que não têm motivação para a participação social, pelos que pertencem a religiões que desprezam o que se passa na vida pública etc.

No presente exercício, à medida que decresce a freqüência com que se acompanha o que acontece na política, as taxa de "não sabe" cresce de 0,8% nos dois primeiro grupos, passa a 1,6% para os que acompanham de vez em quando e saltam abruptamente para os 7,2% entre aqueles que dizem "nunca" acompanhar o que acontece na política.

A Tabela 14, por sua vez, fornece resultados para regiões do país em que o indivíduo residiu até os 16 anos de idade. Consulte-a e verifique que a socialização ocorrida em regiões mais desenvolvidas do país concorre no sentido de enfraquecer a adesão a atitudes como esta de se aceitar a tortura de criminosos como prática legítima corrente.

Tabela 11

Crosstabulation: P62D            CONCORDA TORTURA DE CRIMINOSOS  
 By P20                            CONSTUMA ACOMPANHAR POLITICA

		Count					
P20->	Row Pct	Diaria- mente	1 vez se mana +/-	Veze em quando	Nunca	Row Total	
	Col Pct	1	2	3	4		
	Tot Pct						
P62D							
Concorda	1	41	31	106	46	224	
Total		18.3	13.8	47.3	20.5	28.0	
		17.3	25.8	34.9	33.1		
		5.1	3.9	13.3	5.8		
Depende do Caso	2	47	33	81	34	195	
Total		24.1	16.9	41.5	17.4	24.4	
		19.8	27.5	26.6	24.5		
		5.9	4.1	10.1	4.3		
Discorda	3	147	55	112	49	363	
Total		40.5	15.2	30.9	13.5	45.4	
		62.0	45.8	36.8	35.3		
		18.4	6.9	14.0	6.1		
Nao Sabe	9	2	1	5	10	18	
Total		11.1	5.6	27.8	55.6	2.3	
		.8	.8	1.6	7.2		
		.3	.1	.6	1.3		
Column		237	120	304	139	800	
Total		29.6	15.0	38.0	17.4	100.0	

Number of Missing Observations = 0

Tabela 12

Crosstabulation: P62D            CONCORDA TORTURA DE CRIMINOSOS  
 By P20                    CONSTUMA ACOMPANHAR POLITICA

P20->	Col Pct	Diaria- mente 1	1 vez se- mana +/- 2	Veze em quando 3	Nunca 4	Row Total
P62D						
Concorda Total	1	17.3	25.8	34.9	33.1	224 28.0
Depende do Caso	2	19.8	27.5	26.6	24.5	195 24.4
Discorda Total	3	62.0	45.8	36.8	35.3	363 45.4
Nao Sabe	9	.8	.8	1.6	7.2	18 2.3
Column Total		237 29.6	120 15.0	304 38.0	139 17.4	800 100.0

Number of Missing Observations = 0

Tabela 13

Crosstabulation: P20            CONSTUMA ACOMPANHAR POLITICA  
 By P62D                    CONCORDA TORTURA DE CRIMINOSOS

P62D->	Row Pct	Concorda Total 1	Depende do Caso 2	Discorda Total 3	Nao Sabe 9	Row Total
P20						
Diariamente	1	17.3	19.8	62.0	.8	237 29.6
1 vez semana +/-	2	25.8	27.5	45.8	.8	120 15.0
Veze em quando	3	34.9	26.6	36.8	1.6	304 38.0
Nunca	4	33.1	24.5	35.3	7.2	139 17.4
Column Total		224 28.0	195 24.4	363 45.4	18 2.3	800 100.0

Number of Missing Observations = 0

Tabela 14

Crosstabulation: P62D CONCORDA TORTURA DE CRIMINOSOS  
By P67B REGIAO ONDE VIVEU ATE 16 ANOS

P67B->	Col Pct	Est. S. Paulo 1	Outro do Sudeste 2	Sul 3	Nordeste 4	Outra origem (*) 5	Row Total
P62D							
Concorda Total	1	23.5	29.9	34.5	41.3	(3)	224 28.0
Depende do Caso	2	25.4	23.4	20.7	22.6	(2)	195 24.4
Discorda Total	3	48.9	46.8	44.8	32.3	(6)	363 45.4
Nao Sabe	9	2.3			3.9		18 2.3
Column Total		528 66.0	77 9.6	29 3.6	155 19.4	11 1.4	800 100.0

Number of Missing Observations = 0

(\*). Inclui outras regiões e países. O número total de 11 casos é, mesmo assim, insuficiente para o cálculo de porcentagens. Os números entre parênteses que aparecem nas caselas desta coluna são frequências absolutas.



## 2.5. Apresentação das tabelas de duas entradas

Quando estamos preparando o texto com as conclusões da pesquisa, algumas normas sobre a apresentação de tabelas devem ser seguidas. Usarei a Tabela 15 como modelo.

1) Toda tabela deve ser numerada e receber um título que indique claramente o que ela fornece. No texto, conforme você viu na seção anterior, as tabelas são referidas com maiúsculas, pois são nomes próprios: Tabela 4, Tabela 8 etc. Não escreva "na tabela acima", "na tabela abaixo" etc., pois isto pode complicar a composição do texto.

2) A tabela deve conter apenas os números essenciais. Números em demasia atrapalham a leitura. Inclua as porcentagens de cada casela, usando uma casa decimal (com vírgula, conforme a convenção válida no Brasil). Indique a direção da VI colocando um 100,0% no total de cada uma de suas categorias (sejam elas nas colunas, como está na Tabela 15, ou nas linhas). Sob cada 100,0% escreva o total absoluto de casos naquele grupo. Com este número, o leitor pode recalcular as frequências absolutas das caselas. Além disso, este número de casos indica se há observações suficientes para o cálculo de porcentagens. *Nunca calcule porcentagens para totais inferiores a 25*, pois números pequenos estão sujeitos a flutuações estatísticas muito grandes, podendo distorcer completamente as conclusões.<sup>15</sup>

3) Verifique se a soma das porcentagens, na direção da VI, resulta 100,0% para cada grupo. Por problemas de aproximação, às vezes ao somarmos as taxas podemos encontrar um valor 99,9% ou 100,1%, por exemplo. Você, ao preparar a tabela para edição, deve corrigir estes erros de aproximação subtraindo ou somando a quantia que passou ou sobrou. Faça isto na frequência maior, pois quando o número é grande uma diferença de 0,1 ou 0,2 introduz um erro irrelevante, se comparado com um número já

---

<sup>15</sup> Em caso de totais inferiores a 25, pode-se agregar esta categoria com outra, se isto não prejudicar a análise. Se for conveniente manter em separado, ou quando não se alcançar um número maior que 25 mesmo juntando-se várias categorias pequenas, não calcule porcentagens para este grupo. Na tabela

originalmente pequeno. Compare a Tabela 15, já editada, com a sua original Tabela 9 e veja que foram introduzidas duas aproximações na terceira linha, para evitar totais 99,9%. Quando a frequência absoluta de uma casela é zero, indique isto escrevendo um traço no lugar da porcentagem. Mas se a porcentagem é 0,0 quando a frequência não é nula (exemplo: frequência 1 para o total 2151), escreva 0,0.

4) Use poucas linhas, para não "sujar" a tabela. Necessariamente, use linhas horizontais para abrir e fechar a tabela, e opcionalmente para separar as linhas. Use linhas verticais para separar a coluna de nomenclatura da variável das linhas e para separar a coluna que traz os totais das linhas. Dispense as linhas verticais para separar as caselas. Jamais feche as tabelas nas laterais.

5) No pé da tabela indique sempre a fonte dos dados, pois uma tabela tem certa independência em relação ao texto. Cite instituições, autores, locais, datas, e tudo o que for necessário para deixar bem claro a origem do material. Se foi você mesmo quem fez a pesquisa, cujas especificações por certo foram relatadas na introdução metodológica, basta escrever: "Fonte: Pesquisa de campo."

6) Basicamente, seja claro, econômico e direto.

Com relação ao texto propriamente, é necessário considerar que o leitor nem sempre é capaz de ler uma tabela. Frases do tipo "conforme os dados da Tabela x, pode-se rejeitar a hipótese" não dizem nada. Cite as porcentagens (não precisa citar todas!), chamando a atenção para os achados mais relevantes. No exemplos dados na seção anterior procurei deixar esta prática mais ou menos explícita. Volte aos exemplos, se achar necessário.

Finalmente, em muitas edições você encontrará outras formas de dispor os dados de tabelas de duas entradas, bem como referências feitas no texto. Aqui também a estética e o gosto pessoal têm o seu peso. As revistas científicas costumam trazer as normas que elas seguem na edição dos trabalhos. Na dúvida, seja sempre claro e preciso.

---

mantenha os números absolutos entre parênteses, com uma nota ao pé da tabela dizendo "Total insuficiente para o cálculo de porcentagens". Veja isto ilustrado na Tabela 14.

## MODELO PARA PUBLICAÇÃO

Tabela 15: Posição dos eleitores da Região Metropolitana da Grande São Paulo em relação à tortura de criminosos, segundo renda familiar em número de salários mínimos (SM), 1991

Concordância com a tortura de criminosos	Renda familiar mensal em SM				TOTAL
	1 a 3	4 a 7	8 a 20	21 ou mais	
Concorda totalmente	35,8	32,5	26,2	11,2	28,1
Depende do caso	28,4	22,5	23,4	25,5	24,3
Discorda totalmente	33,8	42,9	48,3	63,3	45,8
Não sabe/não responde	2,0	2,1	2,1	-	1,8
TOTAL	100,0	100,0	100,0	100,0	100,0
% nº de casos	(148)	(240)	(282)	(98)	(768)

Fonte: CEBRAP, Pesquisa sobre Democracia Delegativa