

AVALIAÇÃO DE IMPACTO NA PRÁTICA

GLOSSÁRIO

Amostra aleatória. Também conhecida como “amostra probabilística”. A melhor maneira de evitar uma amostra enviesada ou não-representativa é selecionar uma amostra aleatória. Uma amostra aleatória é definida como aquela na qual a probabilidade de qualquer indivíduo da população ser selecionado é exatamente a mesma de qualquer outro indivíduo.

Amostra. Em estatística, uma amostra é um subconjunto de uma população. Normalmente, a população é muito grande, tornando um censo ou uma enumeração completa de todos os valores da população impraticável ou impossível. A amostra representa um subconjunto de tamanho administrável. As amostras são coletadas e as estatísticas são calculadas de maneira que se possa fazer inferências ou extrapolações da amostra à população. Este processo de coleta de informação a partir de uma amostra é denominada amostragem. (*Wikipédia*).

Atividade. Medidas tomadas ou trabalho realizado através dos quais insumos, tais como fundos, assistência técnica e outros tipos de recursos são mobilizados para produzir resultados específicos.

Atribuição aleatória (Randomized assignment). O método de atribuição aleatória é o mais robusto para estimar contrafactuais e é considerado o padrão-ouro de avaliação de impacto. Com este método, é utilizado um sorteio para decidir quem receberá os benefícios do programa e quem não receberá.

Avaliação de Impacto. Avaliação de impacto procura fazer um nexo de causalidade entre um programa ou intervenção e um conjunto de resultados. Uma avaliação de impacto procura responder se um programa é responsável por mudanças nos resultados de interesse. *Contraste com* Avaliação do processo.

Avaliação do processo. A avaliação do processo procura estabelecer o nível de qualidade ou o sucesso dos processos de um programa: por exemplo, a adequação dos processos administrativos, a aceitabilidade dos benefícios do programa, a clareza da divulgação de informações, a dinâmica interna das organizações de execução, os seus instrumentos de política, seus mecanismos de prestação de serviços, suas práticas de gestão, e as ligações entre estes. *Contraste com:* Avaliação de impacto.

Avaliação. Avaliações são verificações objetivas e periódicas de um projeto, programa ou política planejados, estejam eles em curso ou concluídos. As avaliações são utilizadas para responder questões específicas, geralmente relacionadas com a concepção, implementação e/ou resultados.

Cadeia de Resultados. É a lógica do programa que explica como o objetivo de desenvolvimento deve ser alcançado. Ela mostra a ligação de insumos para atividades para os produtos para os resultados.

Agrupamento (cluster). É um grupo de unidades que são semelhantes, de uma forma ou de outra. Por exemplo, em uma amostra de crianças em idade escolar, as crianças que frequentam a mesma

escola pertencem a um cluster, uma vez que partilham as mesmas instalações da escola, os mesmos professores, e vivem no mesmo bairro.

Comparação antes e depois. Também conhecida como comparação pré-pós ou comparação reflexiva: Uma comparação antes e depois procura estabelecer o impacto de um programa rastreando as alterações nos resultados dos beneficiários do programa ao longo do tempo.

Contrafactual. O contrafactual é uma estimativa do resultado (Y) que um participante do programa teria tido na ausência do programa (P). Por definição, o contrafactual não pode ser observado. Por conseguinte, deve ser estimado.

Correlação “intra-cluster”. É uma correlação (ou similaridade) nos resultados ou características entre as unidades que pertencem ao mesmo cluster. Por exemplo, crianças que frequentam a mesma escola normalmente serão semelhantes ou correlacionados em termos de sua área de residência ou do contexto socioeconômico.

Diferença-em-diferenças. Também conhecida como dupla diferença, D-em-D. Estima o contrafactual para a variação no resultado para o grupo de tratamento, considerando a variação de resultado para o grupo de comparação. Este método permite-nos levar em consideração quaisquer diferenças entre os grupos de tratamento e de comparação que são constantes entre si.

Efeito Hawthorne. O efeito “Hawthorne” ocorre quando o simples facto de que você está observando unidades faz com que estas se comportem de maneira diferente.

Efeito John Henry. O “efeito John Henry” acontece quando o grupo de comparação trabalha mais para compensar o facto de não ter recebido o tratamento. Quando se compara o grupo de tratamento com o grupo “mais trabalhador”, a estimativa do impacto do programa será viesada: iremos estimar um menor impacto do programa do que o verdadeiro impacto que seria encontrado caso o grupo de comparação não fizesse esforço adicional.

Efeito transbordamento (Spillover effect). Também conhecido como contaminação do grupo de comparação. Um efeito de transbordamento ocorre quando o grupo de comparação é afetado pelo tratamento administrado ao grupo de tratamento, embora o tratamento não seja administrado diretamente ao grupo de tratamento. Se os efeitos de transbordamento sobre o grupo de comparação são negativos (ou seja, o grupo é prejudicado por causa do programa), então a diferença direta entre os resultados dos grupos de tratamento e de comparação gerarão uma superestimação do impacto do programa. Em contrapartida, se o efeito de transbordamento no grupo de comparação é positivo (ou seja, gera benefícios), então ele irá produzir uma subestimação do impacto do programa.

Efeito. Mudança, intencional ou não-intencional, devido direta ou indiretamente a uma intervenção.

Estimador. Em estatística, um estimador é uma função da amostra de dados observáveis que é usada para estimar um parâmetro populacional desconhecido, uma **estimativa** é o resultado da aplicação real da função para uma determinada amostra de dados.

Grupo de comparação. Também conhecido como grupo de controle. Um grupo de comparação válido terá as mesmas características que o grupo de beneficiários do programa ("grupo de tratamento"), exceto pelo fato de que o grupo de comparação não se beneficia do programa.

Grupo de Tratamento. Também conhecido como grupo tratado ou grupo de intervenção. O grupo de tratamento é o conjunto de unidades que se beneficiam de uma intervenção.

Hipótese alternativa. Em avaliação de impacto, a hipótese alternativa é geralmente a suposição de que a hipótese nula é falsa; em outras palavras, que a intervenção possui um impacto nos resultados.

Hipótese Nula. A hipótese nula é uma hipótese que pode ser falsificada com base nos dados observados. A hipótese nula tipicamente propõe uma posição geral ou padrão. Em avaliação de impacto, a posição padrão geralmente é que não há diferença entre o grupo de tratamento e o grupo de controle, ou em outras palavras, que a intervenção não possui nenhum impacto nos resultados.

Hipótese: Uma hipótese (do grego *ὑποτιθέναι* – *hypotithenai*) é uma explicação proposta para um fenômeno observável. Veja também: hipótese nula e hipótese alternativa.

Indicador. Um indicador é uma variável que mede um fenômeno de interesse do avaliador. O fenômeno pode ser um insumo, um produto, um resultado, uma característica ou atributo.

Insumos. Os recursos financeiros, humanos e materiais utilizados para a intervenção para o desenvolvimento.

Intenção de tratar ou estimador ITT. O estimador ITT é a diferença direta do indicador do resultado Y para o grupo o qual oferecemos o tratamento e o mesmo indicador para o grupo que não oferecemos o tratamento.

Levantamento de acompanhamento (Follow-up survey). Também conhecido como “pós-intervenção” ou levantamento “ex-post”: um levantamento que é realizado algum tempo após o início do programa, após os beneficiários já tenham se beneficiado do programa por algum tempo. Uma avaliação de impacto pode incluir vários levantamentos de acompanhamento.

Linha de base. Pré-intervenção, ex-ante. A situação anterior à intervenção, contra a qual o progresso pode ser avaliado ou comparações serem feitas.

Matching. “Matching” é um método de avaliação não-experimental que utiliza grandes conjuntos de dados e pesadas técnicas estatísticas para construir o melhor grupo de comparação “artificial” possível para um determinado grupo de tratamento.

Métodos de seleção aleatória. Método de seleção aleatória é o nome do grupo formado por vários métodos que utilizam a seleção aleatória para identificar o contrafactual. Entre eles estão: a atribuição aleatória do tratamento, a oferta aleatória do tratamento e a promoção aleatória.

Modelo de Regressão por descontinuidade (RDD). O Modelo de Regressão por descontinuidade (RDD) é um método de avaliação não-experimental. É adequado para programas

que utilizam um índice contínuo para classificar os potenciais beneficiários e que possuem um ponto divisório ao longo do índice que determina se os potenciais beneficiários do programa receberão ou não o benefício.

Monitoramento. Monitoramento é um processo contínuo de recolhimento e análise de informações para verificar quão bem um projecto, programa ou política, está sendo realizado. Baseia-se principalmente em dados administrativos para monitorar o desempenho contra os resultados esperados, fazer comparações entre programas e analisar as tendências ao longo do tempo. Normalmente o monitoramento acompanha insumos, atividades e produtos, embora ocasionalmente também inclua resultados. O monitoramento é utilizado para informar o dia-a-dia da gestão e decisões tomadas.

Nível de significância. O nível de significância é usualmente indicado pelo símbolo grego α (alfa). Níveis comuns de significância são de 5% (0,05), 1% (0,01) e 0,1% (0,001). Se um teste de significância dá um p-valor menor que o nível α , a hipótese nula é rejeitada. Tais resultados são informalmente referidos como "estatisticamente significantes". Quanto menor o nível de significância, mais forte são as evidências geradas. Escolher o nível de significância é uma tarefa arbitrária, mas para muitas aplicações, um nível de 5% é escolhido, por nenhuma outra razão além do fato de ser convencional.

Oferta aleatória. Oferta aleatória é um método para identificar o impacto de uma intervenção. Quando o administrador do programa pode escolher aleatoriamente as unidades a quem irá oferecer tratamento a partir do universo de unidades elegíveis, mas não pode obter o cumprimento perfeito: não é possível forçar qualquer unidade a participar/aceitar o tratamento, e não é possível recusar uma unidade a participar, se a mesma insiste em participar/ser tratada. Neste método, a oferta aleatória do programa é utilizada como variável instrumental para a real participação no programa.

Poder estatístico. O poder de um teste estatístico é a probabilidade de que o teste irá rejeitar a hipótese nula quando a hipótese alternativa é verdadeira (isto é, que não vai cometer um erro de Tipo II). À medida que aumenta o poder, as chances de ocorrer o erro do tipo II diminuem. A probabilidade de um erro do tipo II é referida como a falsa taxa negativa (β). Portanto, o poder é igual a $1 - \beta$.

Produto (Output). Os produtos, bens de capital e serviços que são produzidos (fornecidos) diretamente por uma intervenção.

Promoção aleatória. Promoção aleatória é um método similar ao de oferta aleatória. Em vez de selecionar aleatoriamente a quem o tratamento será oferecido, selecionamos aleatoriamente a quem o tratamento será promovido. Desta forma, podemos deixar o programa aberto a todos.

Regressão. Em estatística, a análise de regressão inclui todas as técnicas de modelagem e análise de diversas variáveis, quando o foco está no relacionamento entre uma variável dependente e uma ou mais variáveis independentes. Na avaliação de impacto, a análise de regressão nos ajuda a compreender como o valor típico do indicador de resultado Y (variável dependente) muda quando o valor do tratamento ou do grupo de comparação P (variável independente) é variado, enquanto as características dos beneficiários (outras variáveis independentes) são mantidas fixas.

Resultado. Pode ser intermediário ou final. O resultado é uma consequência de interesse que surge através de uma combinação de fatores de oferta e demanda. Por exemplo, se uma intervenção conduz a uma maior oferta de serviços de vacinação, então os números reais de vacinação seria um resultado. Contudo, estes números não dependem apenas do fornecimento de vacinas, mas também do comportamento dos beneficiários: eles vão ao posto para serem vacinados? Resultados finais/de longo prazo são resultados distantes; essa distância pode ser interpretada por uma dimensão de tempo (leva um longo tempo para se chegar ao resultado), ou uma dimensão causal (são necessários muitos links para se atingir o resultado).

Tratamento nos tratados (Efeito do). Também conhecido como TOT. O efeito do tratamento sobre os tratados é o impacto do tratamento sobre as unidades que realmente se beneficiaram com o tratamento. *Contraste com:* Intenção de tratar.

Universo. O grupo de unidades que são elegíveis para receber uma intervenção ou tratamento.

Validade externa. Validade externa significa que o impacto causal descoberto na avaliação de impacto pode ser generalizado para o universo de todas as unidades elegíveis. Para uma avaliação possuir validade externa, é necessário que sua amostra seja representativa do universo.

Validade interna. Validade interna significa que a avaliação do impacto usa um grupo de comparação válido, ou seja, um grupo de comparação que é uma estimativa válida do contrafactual.

Variável instrumental. A variável instrumental é uma variável que ajuda a identificar o impacto causal de um programa, quando a participação no mesmo é parcialmente determinada pelos potenciais beneficiários. Uma variável deve possuir duas características para ser classificada como boa variável instrumental: (i) deve ser correlacionada com a participação no programa, e (ii) não pode ser correlacionada com os resultados (Y) (para além de através da participação do programa) ou com variáveis não observáveis.

Variável. Uma variável é um símbolo que representa um valor que pode alterar.

Viés de seleção. Viés de seleção ocorre quando as razões pelas quais um indivíduo participa de um programa são correlacionadas com os resultados. Este viés ocorre geralmente quando o grupo de comparação é inelegível ou se auto-seleciona para fora do tratamento.

Viés. O viés de um estimador é a diferença entre a expectativa do estimador e o verdadeiro valor do parâmetro a ser estimado. Na avaliação de impacto, esta é a diferença entre o impacto que você calcula e o impacto real do programa.