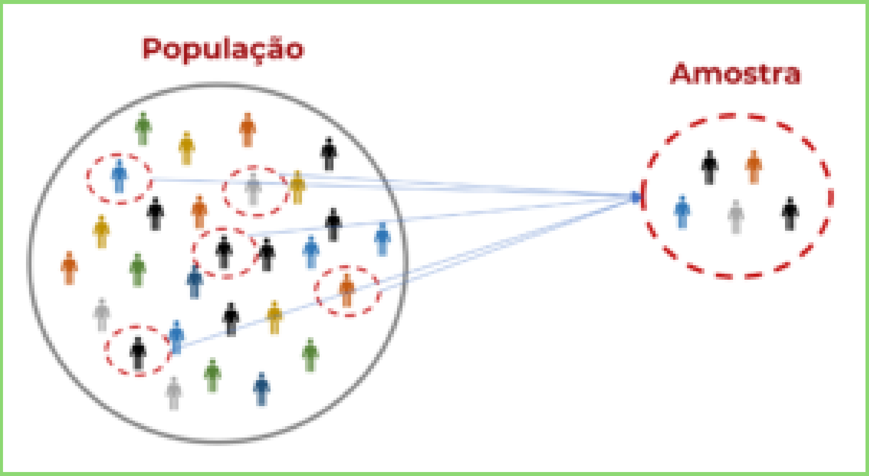
Elirez Bezerra da Silva

Amostragem



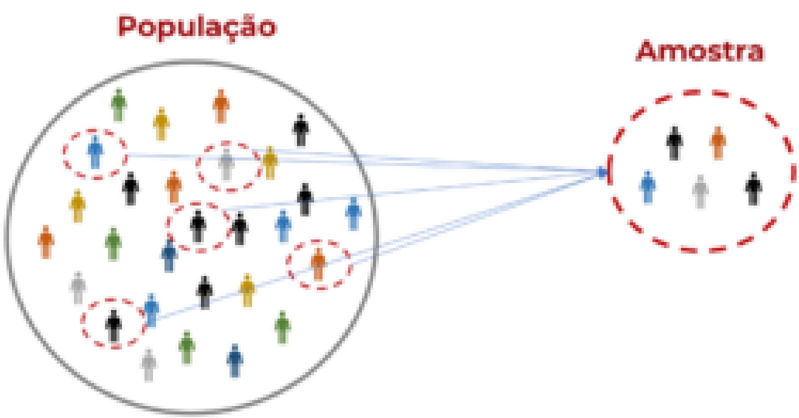
Pedro & João editores

Amostragem

Pedro & João editores
20 Anos

Elirez Bezerra da Silva

Amostragem



Pedro & João editores
20 Anos

Elirez Bezerra da Silva

Amostragem. São Carlos: Pedro & João Editores, 2025. 103p. 16 x 23 cm.

ISBN: 978-65-265-1880-9 [Digital]

1. Amostra. 2. Cálculo. 3. Proporção. 4. Pesquisa. I. Título.

CDD – 370

Capa e diagramação: Elirez Bezerra da Silva
Ficha Catalográfica: Hélio Márcio Pajeú – CRB - 8-8828
Editores: Pedro Amaro de Moura Brito & João Rodrigo de Moura Brito

Pedro & João Editores
www.pedroejoaoeditores.com.br
13568-878 – São Carlos – SP
2025

Agradecimentos

Ao Exército Brasileiro por ter me autorizado a fazer os cursos de mestrado e doutorado.

A Escola de Educação Física do Exército pela minha formação em educação física e por ter me proporcionado praticar ensino, pesquisa e desporto por quase 20 anos.

Ao Instituto de Pesquisa da Capacitação Física do Exército pela oportunidade de desenvolver pesquisas na área da aptidão física e saúde.

Ao Instituto Brasileiro de Medicina e Reabilitação pela minha formação em fisioterapia.

A Universidade Gama Filho (UGF) por ter me acolhido para os cursos de mestrado e doutorado.

Ao Prof. Dr. Paulo Sérgio Chagas Gomes por ter sido meu orientador nos dois cursos e ter aberto a janela para o mundo científico.

Ao Programa de Pós-graduação em Ciências do Exercício e Esporte da Universidade do Estado do Rio de Janeiro (PPGCEE / UERJ) pelo meu trabalho como professor visitante e professor pesquisador.

Ao Programa de Mestrado Profissional em Gestão do Trabalho para a Qualidade do Ambiente Construído da Universidade Santa Úrsula pelo meu trabalho como professor permanente.

Aos meus alunos de mestrado, doutorado e integrantes do Grupo de Pesquisa em Ciências do Exercício e da Saúde (www.gpces.com.br) pela oportunidade em ensinar e aprender.

A Guillermo Brito Portugal pela revisão.

Palavras do autor

Certa vez, escutando o rádio do carro enquanto dirigia para o meu trabalho, uma repórter entrevistando um pesquisador brasileiro, que acabara de ser agraciado com a Medalha Fields, pediu para que ele respondesse sobre um assunto com base na ciência. Após a resposta, a repórter disse “agora os nossos ouvintes e eu temos a certeza sobre este assunto, já que a sua resposta se baseou na ciência”. Prontamente, o pesquisador replicou “mas quem lhe disse que a ciência dá certeza? Nós cientistas, a única certeza que temos é do quanto estamos errados. Claro que trabalhamos para que este erro seja mínimo, seja pequeno”. Sábias palavras desse pesquisador. A ciência é cercada por diversas fontes de erro, cabendo ao bom pesquisador identificá-las e neutralizá-las, mas nem sempre isto é possível.

A principal fonte de erro na ciência é a amostragem, que é o ato de quantificar e caracterizar a amostra necessária para participar de uma pesquisa. Para controlar esta principal fonte de erro em pesquisas, este livro orienta como calcular o tamanho de uma amostra para os seguintes desenhos das pesquisas: (1) de um grupo e duas medidas repetidas; (2) de dois grupos e uma medida; (3) de três ou mais grupos e uma medida; (4) de um grupo e três ou mais medidas repetidas; (5) de dois ou mais grupos e duas ou mais medidas repetidas; (6) de correlação entre duas variáveis; (7) de predição de uma variável contínua intervalar a partir de duas ou mais variáveis; (8) de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis; (9) de predição de uma variável nominal, que expressa uma quantidade, a partir de duas ou mais variáveis, utilizando o programa gratuito G*Power 3.1.9.7, escrito por Franz Faul, da Universidade de Kiel / Alemanha, disponível em https://www.kent.ac.uk/software/gpower . Para cada desenho da pesquisa acima, os campos do G*Power 3.1.9.7 para a entrada dos dados estão em ***negrito*** e ***itálico***, para facilitar o seu manuseio.

O livro mostra ainda como calcular o tamanho da amostra para determinar a média ou proporção de uma variável da população, como alocar os participantes nos grupos de pesquisa e como agir quando não se alcança o tamanho amostral desejado.

A você, desejo uma boa leitura.

Elirez Bezerra da Silva

Acesso ao currículo: http://lattes.cnpq.br/7486340493431857

Prefácio

Conheço o Prof. Dr. Elirez Bezerra da Silva há mais de 30 anos. Nos anos 1990, ele atuava na Escola de Educação Física do Exército, e tivemos a oportunidade de interagir quando ministrei cursos de pós-graduação oferecidos por aquela instituição. Foi, portanto, com grande alegria que recebi seu convite para prefaciar esta obra.

Tive a oportunidade de acompanhar a trajetória acadêmica e profissional do Prof. Elirez, cuja formação sólida e experiência nas áreas de Educação Física e Fisioterapia são reconhecidas por seus pares. Graduou-se em Educação Física pela Escola de Educação Física do Exército em 1984 e, posteriormente, em Fisioterapia pelo Instituto Brasileiro de Medicina de Reabilitação em 1997. Buscando aprofundar seus estudos, obteve o título de mestre em Educação Física pela Universidade Gama Filho em 1999 e concluiu o doutorado na mesma instituição em 2004. Aliás, tive a honra de compor o júri que avaliou sua tese.

Atualmente, o Prof. Elirez é docente do Programa de Pós-Graduação em Ciências do Esporte e do Exercício da Universidade do Estado do Rio de Janeiro, onde contribui significativamente para o campo do exercício físico e da saúde. Ele ministra regularmente a disciplina de Estatística Aplicada às Ciências do Exercício e do Esporte, trazendo para a sala de aula sua ampla experiência com o gerenciamento de projetos de pesquisa e bancos de dados. Sua produção acadêmica reflete essa expertise, com publicações que abordam desde técnicas estatísticas até meta-análises desenvolvidas por seu grupo de estudos. Nesse contexto, parece-me que o livro "Amostragem" nasce como consequência natural da experiência profissional e acadêmica angariada pelo autor.

A ciência, por sua própria essência, busca compreender o mundo por meio da formulação de perguntas e da investigação sistemática. Nesse processo, a mensuração e análise de dados desempenham um papel central e a amostragem se estabelece como um dos pilares fundamentais para garantir a representatividade dos resultados obtidos. Ao levar isso em consideração, esta obra se apresenta como uma contribuição essencial para pesquisadores, professores e estudantes interessados na análise quantitativa dos fenômenos. Em um campo no qual a precisão dos dados é crucial, compreender os diferentes tipos de amostragem, os cálculos estatísticos envolvidos e as estratégias para minimizar erros torna-se indispensável para a construção de um conhecimento sólido.

A abordagem adotada pelo autor vai além da simples descrição dos fundamentos da amostragem, pois também fornece um guia prático para a aplicação de métodos estatísticos em diversos contextos de pesquisa. A compreensão desses processos é um diferencial para a tomada de decisões baseadas em evidências, seja na realização de estudos epidemiológicos, na avaliação de programas educacionais ou no desenvolvimento de políticas públicas. Assim, cada capítulo do livro é estruturado como uma etapa fundamental para a compreensão da relação entre a amostra estudada e a população de interesse. O autor conduz o leitor de maneira progressiva e didática, desde a definição dos critérios de seleção até a análise dos fatores determinantes para o cálculo amostral, proporcionando uma compreensão clara e aplicada do tema.

Um dos grandes destaques é o detalhamento do uso do programa G*Power 3.1.9.7, uma ferramenta de análise estatística amplamente utilizada para calcular o tamanho adequado de amostras em diferentes desenhos de estudo. O autor oferece um passo a passo detalhado para sua utilização, ensinando como inserir dados corretamente e interpretar os resultados obtidos. Tal abordagem se revela de grande utilidade, visto que o G*Power é um programa de domínio público frequentemente utilizado por estudantes, que muitas vezes enfrentam dificuldades tanto no manuseio da ferramenta quanto na interpretação dos resultados.

Nos capítulos seguintes, o autor explora fatores determinantes do cálculo amostral, como erro alfa, poder estatístico e tamanho do efeito. Esses elementos são apresentados de forma didática, com exemplos práticos que facilitam a compreensão de como afetam a precisão dos resultados. Além disso, a obra demonstra como diferentes tipos de análise demandam tamanhos de amostra específicos para garantir a confiabilidade dos achados. Outro ponto relevante abordado é a questão da insuficiência amostral, que ocorre quando o tamanho da amostra estimado inicialmente não é alcançado ao final do estudo, seja por dificuldades na coleta de dados ou pela desistência de participantes.

Para lidar com essa limitação, o livro propõe estratégias eficazes, incluindo o recálculo do poder estatístico e a avaliação do impacto da redução da amostra na validade dos resultados. Além disso, a obra se destaca pela ênfase na interpretação correta dos resultados estatísticos, uma vez que muitos estudos sofrem com equívocos na análise, o que pode levar a conclusões errôneas. O autor reforça, portanto, a importância não apenas de calcular corretamente o tamanho da amostra, mas também de compreender os pressupostos estatísticos subjacentes a cada método de análise, garantindo maior rigor e confiabilidade na pesquisa científica.

Ao final da leitura, fica evidente que a amostragem é muito mais do que um mero cálculo matemático; ela constitui um componente essencial da credibilidade científica. Como bem destaca o autor, "a ciência é cercada por diversas fontes de erro, cabendo ao bom pesquisador identificá-las e neutralizá-las". Nesse sentido, "Amostragem" se apresenta como uma ferramenta valiosa para aqueles que desejam conduzir pesquisas com rigor metodológico e solidez analítica.

Parabenizo o Prof. Elirez por esta obra essencial e agradeço pela honra de prefaciá-la. A riqueza das informações aqui contidas certamente ajudará os leitores não apenas em seus projetos pontuais, mas também em sua formação científica mais ampla. Que esta leitura seja proveitosa e inspiradora.

Boa leitura!

Prof. Dr. Paulo T.V. Farinatti

Professor Titular

Universidade do Estado do Rio de Janeiro

Acesso ao currículo: http://lattes.cnpq.br/6076797913133653

Sumário

Amostragem

Definição e tipos

Definir amostragem é fácil. Difícil é fazê-la com o mínimo de erro. Toda seleção, extração ou formação de um grupo de participantes para uma pesquisa é uma amostragem. Entretanto, se feita desta maneira, a conclusão desta pesquisa estará errada. Sendo fundamental para diminuir o tempo, o custo e o trabalho durante a execução de uma pesquisa, pode se transforar na principal fonte de erro da pesquisa se não for feita adequadamente.

A amostragem pode ser aleatória, chamada também de randômica, ou por conveniência. A amostragem aleatória é probabilística, enquanto a amostragem por conveniência é não-probabilística. Na amostragem aleatória cada participante tem a mesma probabilidade de participar da pesquisa, enquanto na amostragem por conveniência os participantes da pesquisa são escolhidos de maneira mais fácil ou acessível para o pesquisador. Você deve estar se perguntando qual delas devo usar? Vai depender do desenho da pesquisa. Para pesquisas experimentais verdadeiras, aquelas que comparam grupos e controlam o experimento, irremediavelmente você terá que usar a amostragem aleatória. Por quê?

A resposta começa lá no início do livro nas "Palavras do autor": "A ciência é cercada por diversas fontes de erro, cabendo ao bom pesquisador identificá-las e neutralizá-las, mas nem sempre isto é possível". Estas diversas fontes de erros são as diferenças entre os participantes da pesquisa. Por exemplo, sexo, idade, nível de aptidão e atividade física, nível educacional, nível social, cultura, humor, doenças prévias, doenças não diagnosticadas, nível de estresse, estado de saúde, profissão, horas de sono, quantidade e qualidade alimentar etc. são diferentes entre os participantes da pesquisa. Todas estas fontes de erro e mais aquelas englobadas pelo “etc...” são variáveis de confusão que interferirão no resultado da pesquisa, “cabendo ao bom pesquisador identificá-las e neutralizá-las”. Porém, como neutralizar tantas variáveis assim?

A neutralização destas variáveis de confusão ocorre quando a seleção dos participantes é aleatória. Como cada participante tem a mesma probabilidade de participar da pesquisa, o acaso vem em socorro ao pesquisador e faz o contrabalanço de todas estas variáveis de confusão, tornando os grupos homogêneos.

Os tipos de amostragem aleatória são simples, estratificada, sistemática e por conglomerados.

A amostragem aleatória simples é um método imparcial e probabilístico, que seleciona uma amostra representativa de uma população, na qual cada participante da população tem a mesma probabilidade de ser selecionado para a amostra.

A amostragem aleatória estratificada é uma amostragem aleatória simples que considera as várias características (estratos) de uma população. Em outras palavras, o percentual destes estratos na amostra deve ser o mesmo percentual destes estratos na população.

A amostragem sistemática é uma amostragem probabilística que faz a seleção aleatória do primeiro participante e os participantes seguintes são selecionados obedecendo intervalos fixos ou sistemáticos até chegar ao tamanho da amostra estimado. A meu ver, este tipo de amostragem fere frontalmente o princípio “cada participante tem a mesma probabilidade de participar da pesquisa”.

A amostragem por conglomerados divide a população em grupos, conglomerados ou clusters com base em sua localização geográfica ou organizacional, para em seguida, fazer uma amostragem aleatória simples dentro de cada grupo, conglomerado ou clusters. Ela se assemelha à amostragem aleatória estratificada, mas esta considera os “estratos”, que são as características dos participantes, enquanto por conglomerados considera a “localização geográfica ou organizacional” dos participantes.

Se por um lado a amostragem aleatória simples é a que mais neutraliza as variáveis de confusão ou as diversas fontes de erros na pesquisa, por outro lado ela pode criar outra fonte de erro que é a desigualdade da quantidade de participantes em cada grupo, principalmente quando o tamanho da amostra estimado for pequeno.

No intuito de neutralizar esta fonte de erro, surgiu a amostragem aleatória por minimização que consiste em uma amostragem aleatória simples até que um dos grupos alcance a quantidade de igualdade entre os grupos. A partir daí, os participantes remanescentes são alocados aleatoriamente para os outros grupos, se três ou mais grupos, ou alocados por conveniência para o outro grupo, se dois grupos. Um exemplo para este texto ficar mais claro: suponhamos que o tamanho da amostra seja de 45 participantes para serem alocados aleatoriamente para os grupos A, B e C. A quantidade de igualdade entre os grupos será de 45/3 = 15 participantes. Em dado momento da alocação o grupo A tem 15 participantes e os grupos B e C têm 3 e 4 participantes, respectivamente. Os outros 23 participantes remanescentes serão alocados entre os grupos B e C. Em dado momento da alocação o grupo C tem 15 participantes e o grupo B tem 12 participantes. Os outros 3 participantes remanescentes serão alocados para o grupo B.

Diferentemente da amostragem aleatória, na qual o acaso determina a seleção dos participantes, a amostragem por conveniência é não-probabilística, sendo a seleção dos participantes determinada pela conveniência ao pesquisador. Como o acaso não faz parte dela, ela é singular e se aplica nos desenhos de pesquisas pragmáticas e qualitativas.

Fatores determinantes do cálculo amostral para testar a diferença de médias ou proporções entre grupos ou testar a relação entre duas ou mais variáveis

Para que a estimativa da quantidade de participantes de uma pesquisa tenha o mínimo de erro é indispensável que o cálculo amostral considere como fatores determinantes: o desenho da pesquisa, o teste estatístico a ser utilizado, o erro α, o poder (1 – erro β) e o tamanho do efeito.

Se considerados os fatores determinantes do cálculo amostral: desenho da pesquisa e teste estatístico; erro α e poder (1 – erro β); e tamanho do efeito, então esta quantidade amostral poderá ser utilizada para (1) testar a diferença de médias ou proporções entre grupos ou (2) testar a relação entre duas ou mais variáveis.

Desenho da pesquisa e teste estatístico

As principais características do desenho da pesquisa consideradas para o cálculo amostral são as quantidades de grupos, de medidas repetidas do desfecho e de variáveis independentes. Consequentemente, o teste estatístico correspondente para tal desenho deve ser utilizado na pesquisa e selecionado para o cálculo amostral.

Erro α e poder (1 – erro β)

Erro α e poder (1 – erro β) são duas probabilidades distintas de afirmar que o resultado da pesquisa não foi devido ao acaso, ou seja, foi significativo. O erro α é a probabilidade de o pesquisador estar errado ao fazer tal afirmação. Por isto este erro deve ser pequeno, no máximo até 5%. Já o poder (1 – erro β) é a probabilidade de o pesquisador estar certo ao fazer tal afirmação. Por isto este acerto deve ser grande, no mínimo de 80%. Normalmente, são considerados o erro máximo e o acerto mínimo, ou seja, α igual a 0.05 e poder (1 – erro β) igual a 0.80 para o cálculo amostral.

Tamanho do efeito

O tamanho do efeito, como o próprio nome diz, é a diferença entre os resultados de dois grupos. Se considerarmos que cada grupo terá como resultado uma média ± desvio padrão ou uma proporção, então o tamanho do efeito será a diferença entre as médias ou proporções.

O fator determinante que mais merece atenção durante o cálculo amostral é o tamanho do efeito. A razão para isto é simples: há diversos tamanhos dos efeitos. O porquê de haver diversos tamanhos dos efeitos é compreensível. Para poder comparar os tamanhos dos efeitos de diferentes desfechos, a estatística padronizou os tamanhos dos efeitos. A razão entre a diferença de médias e a variabilidade dos resultados foi a “mágica” utilizada pela estatística para padronizar os tamanhos dos efeitos e, com isto, poder comparar os tamanhos dos efeitos de diferentes desfechos. Se não fizesse isto, um tamanho do efeito grande para um determinado desfecho poderia ser moderado ou pequeno para outro desfecho, ou ainda, seria impossível comparar tamanhos dos efeitos de diferentes unidades de medidas.

A diversidade de tamanhos dos efeitos decorre das diferentes maneiras de se calcular a variabilidade, que é o denominador da razão acima. Dependendo do teste estatístico, o denominador poderá utilizar o desvio padrão de um ou mais grupos, a variância de um ou mais grupos, a correlação e o coeficiente de determinação entre as medidas repetidas.

Esta diversidade é muito bem evidenciada no aplicativo G*Power 3.1.9.7, que utiliza 11 diferentes tipos de tamanhos dos efeitos (vide Tabela 1) para diferentes finalidades, como: (1) calcular o tamanho da amostra estimado, sendo dados o erro α, poder (1 – erro β) e tamanho do efeito; ou (2) calcular o erro α, sendo dados o poder (1 – erro β), tamanho do efeito e tamanho da amostra; ou (3) calcular o poder (1 – erro β) alcançado, sendo dados o erro α, tamanho do efeito e tamanho da amostra; ou (4) calcular o tamanho do efeito, sendo dados o erro α, poder (1 – erro β) e tamanho da amostra. Mais à frente, mostrarei como são calculados os tamanhos dos efeitos utilizados no cálculo amostral dos desenhos de pesquisas propostos neste livro.

Tabela 1 - Onze diferentes tipos de tamanhos dos efeitos exigidos pelo G*Power 3.1.9.7 para o cálculo amostral

Tamanho do efeito	Família dos testes t *		Família dos testes F **		Família dos testes z	Família dos testes χ^2	Família dos testes exatos	
Pequeno	ρ = .10	d = .20	f = .10	f^2 = .02	q = .10	w = .10	g = .05	ρ = .10
Médio	ρ = .30	d = .50	f = .25	f^2 = .15	q = .30	w = .30	g = .15	ρ = .30
Grande	ρ = .50	d = .80	f = .40	f^2 = .35	q = .50	w = .50	g = .25	ρ = .50

* Há ainda o tamanho do efeito dz, para o qual não há classificação

** Há ainda os tamanhos de efeitos Δ, f^2(V) e f(V), para os quais não há classificação

A determinação de um dos fatores determinantes, seja o tamanho da amostra ou o erro α ou o poder (1 – erro β) ou o tamanho do efeito, requer a interação com os outros fatores determinantes remanescentes. Um menor tamanho da amostra é requerido quando o tamanho do efeito for maior, o erro α for maior, o poder (1 – erro β) for menor, a quantidade de grupos for menor, a quantidade de medidas repetidas for maior e a correlação entre as medidas repetidas for maior (vide Quadro 1).

Quadro 1 – Tamanho da amostra estimado em função dos fatores determinantes

<table>
<tr><td>Fatores determinantes</td><td>Cálculo amostral</td></tr>
<tr><td>Tamanho do efeito grande</td><td rowspan="3">Tamanho da amostra será menor</td></tr>
<tr><td>Erro α maior</td></tr>
<tr><td>Poder (1 – erro β) menor</td></tr>
</table>

A menor quantidade de grupos, a maior quantidade de medidas repetidas e a maior correlação entre as medidas repetidas requer também um tamanho de amostra menor

Determinação do tamanho do efeito para o cálculo amostral

Considerando a dificuldade que se tem para recrutar participantes para uma pesquisa, seria razoável utilizar o tamanho do efeito grande (vide Tabela 1) para que o tamanho da amostra fosse menor (vide Quadro 1). Entretanto, esta não é a melhor decisão, porque o tamanho do efeito verdadeiro encontrado ao final da pesquisa poderá ser menor que o tamanho do efeito grande utilizado para estimar o tamanho da amostra menor. Consequentemente, para manter o erro α no máximo de 0.05 e o poder (1 – erro β) no mínimo de 0.80, o tamanho do efeito verdadeiro menor obtido ao final da pesquisa irá requerer uma amostra maior do que aquela estimada pelo cálculo amostral para pesquisa. Então, utilizar o tamanho do efeito grande não é o melhor caminho para a estimação do tamanho da amostra.

A princípio, o mais razoável seria utilizar o tamanho do efeito pequeno, que consequentemente vai requerer um tamanho da amostra maior. Mas, uma amostra maior não iria dificultar ainda mais o recrutamento de participantes para uma pesquisa? Sim, é verdade. Porém, o tamanho do efeito verdadeiro encontrado ao final da pesquisa poderá ser maior que o tamanho do efeito pequeno utilizado para estimar o tamanho da amostra. Por conseguinte, o poder (1 – erro β) estimado em 0.80 para o cálculo amostral poderá ser maior, aumentando a confiança para afirmar que o resultado da pesquisa não foi devido ao acaso, ou seja, foi significativo. Outra razão, não menos importante, para ser utilizado um tamanho do efeito pequeno é que, apesar da dificuldade que se tem para recrutar participantes para uma pesquisa, não se pode descartar a possibilidade de recrutamento deste tamanho amostral estimado.

Há três formas para se determinar o tamanho do efeito: (1) arbitrariamente; ou (2) a partir dos resultados de um estudo piloto; ou (3) a partir dos resultados de uma pesquisa semelhante publicada. Minha sugestão é fazer o cálculo amostral utilizando tamanho do efeito "pequeno" previsto na Tabela 1 e ao final da pesquisa, caso o tamanho da amostra estimado não seja alcançado na pesquisa, calcular o poder (1 – erro β), ou seja, analisar os dados considerando a hipótese nula (H_0) falsa, utilizando o tamanho do efeito obtido na pesquisa. Mais à frente, explicarei melhor isto em "Nem sempre o tamanho da amostra estimado é alcançado ao final da pesquisa: o que fazer?" na p. 58.

Cálculo amostral para testar a diferença de médias entre grupos

Diferentes desenhos de pesquisar podem surgir da diferença de médias entre grupos, que dependerá da quantidade de grupos e da quantidade de medidas repetidas: (1) de um grupo com duas medidas repetidas; (2) de dois grupos com uma medida; (3) de três ou mais grupos com uma medida; (4) de um grupo com três ou mais medidas repetidas; (5) de dois ou mais grupos com duas ou mais medidas repetidas.

Tamanho da amostra para uma pesquisa de um grupo com duas medidas repetidas

O desenho desta pesquisa se caracteriza por uma amostra selecionada aleatoriamente para executar um determinado experimento X, cujo desfecho Y será avaliado em dois momentos Y_1 e Y_2. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Tail(s); α err prob; Effect size dz; e Power (1 – β prob).*** Preencha estes campos com os dados abaixo e vide Figura 1:

> ***Test Family*** *= t tests*
>
> A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.
>
> ***Statistical test*** *= Means: Difference between two dependente means (matched pairs)*

O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.

Type of power analysis *= A priori: Compute required sample size – given α, power, and effect size*

O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.

Tail(s) *= Two ou One*

Two (duas caudas) devem ser selecionadas quando a diferença de médias pode ser tanto positiva como negativa. *One* (uma cauda) deve ser selecionada quando a diferença de médias pode ser somente positiva ou negativa.

A err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a diferença entre as médias das medidas repetidas não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Effect size dz *=?*

O tamanho do efeito pequeno é o recomendável para o cálculo amostral (vide Tabela 1). Porém, o tamanho do efeito dz ainda não tem classificação. A solução para isto será apresentada mais à frente.

Power (1 – β prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a diferença entre as médias das medidas repetidas não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral (vide Figura 1).

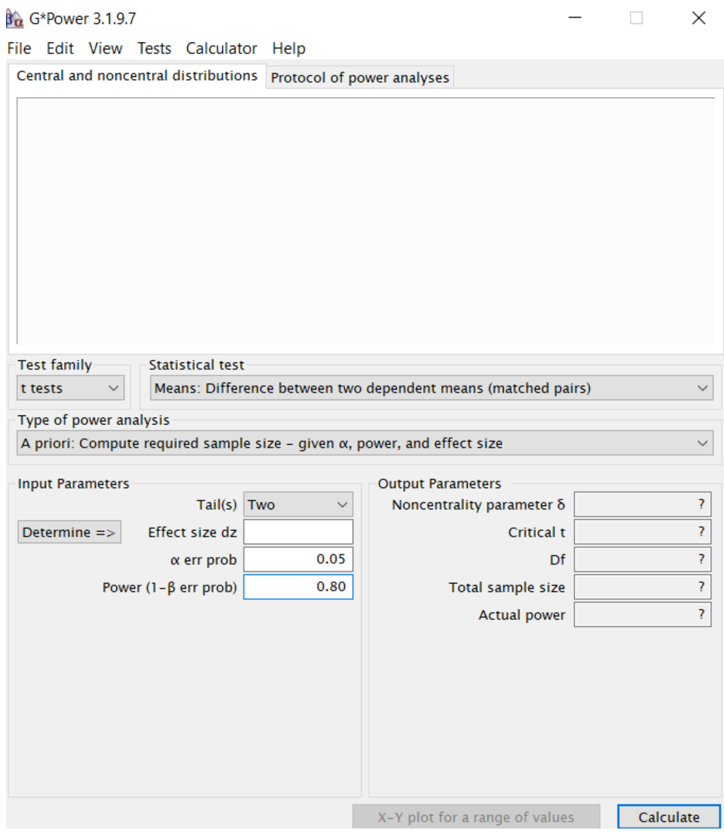


Figura 1 – Cálculo amostral para uma pesquisa de um grupo com duas medidas repetidas. O tamanho da amostra (***Total sample size***) não foi calculado porque não há uma classificação para o tamanho do efeito dz (Tabela 1)

Como não há uma classificação para o tamanho do efeito dz, ou seja, não sabemos o tamanho do efeito dz "pequeno", então ele deve ser calculado a partir dos resultados de um estudo piloto ou de uma pesquisa semelhante publicada.

Supondo que os resultados de uma pesquisa semelhante publicada tenham sido média pré-intervenção = 20; média pós-intervenção = 15; desvio padrão pré-intervenção = 7; desvio padrão pós-intervenção = 5; correlação entre as medidas pré e pós = 0.5.

Para determinar o tamanho do efeito dz, clique em ***Determine*** = >, que abrirá um quadro secundário. Selecione ***From group parameters***. Nos campos ***Mean group 1*** informe a média pré-intervenção = 20; ***Mean group 2*** informe a média pós-intervenção = 15; ***SD group 1*** informe o desvio padrão pré-intervenção = 7; ***SD group 2*** informe o desvio padrão pós-intervenção = 5; ***Correlation between groups*** informe a correlação entre as medidas pré-intervenção e pós-intervenção = 0.5. Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito dz calculado no quadro secundário e igual a 0.8006408 foi transferido automaticamente para o campo ***Effect size dz*** do quadro principal. Neste quadro, clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 15 participantes (vide Figura 2). O G*Power calculou o tamanho do efeito dz pela fórmula abaixo:

$$\text{Tamanho do efeito dz} = \frac{(\text{Média pós-intervenção} - \text{Média pré-intervenção})}{\text{Raiz (DP pós-intervenção}^2 + \text{DP pré-intervenção}^2 - 2 * r * \text{DP pós-intervenção} * \text{DP pré-intervenção)}}$$

sendo DP = desvio padrão e r = coeficiente de correlação de Pearson entre as medidas pré-intervenção e pós-intervenção.

O outro caminho para determinar o tamanho do efeito dz no quadro secundário é ***From diferences***. Neste caminho, inserir no campo ***Mean of difference*** a diferença entre as médias pós-intervenção e pré-intervenção e inserir no campo ***SD of difference*** o erro padrão da diferença. Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito dz calculado no quadro secundário será transferido automaticamente para o campo ***Effect size dz*** do quadro principal.

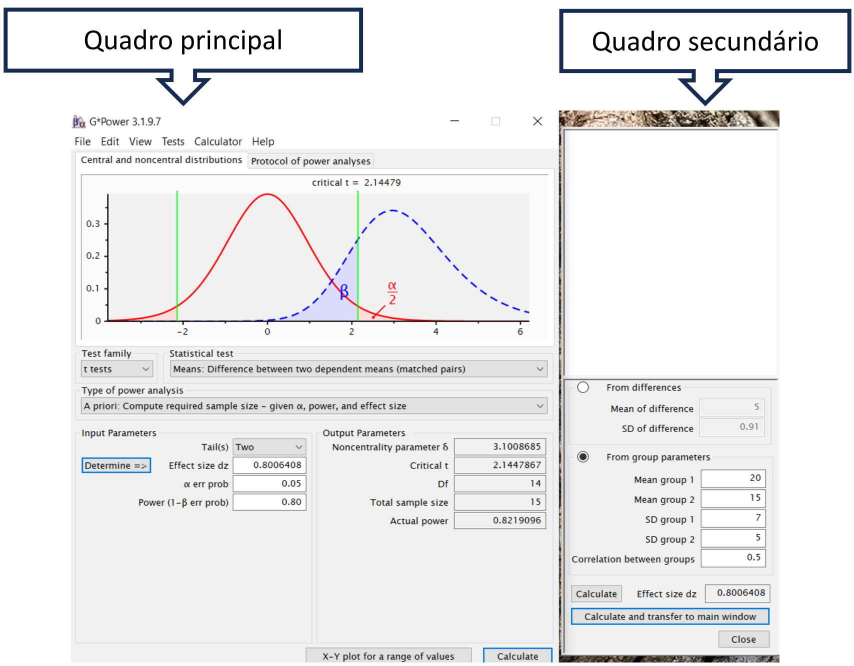


Figura 2 – Determinação do tamanho do efeito dz a partir de resultados de uma pesquisa semelhante publicada (quadro secundário)

Tamanho da amostra para uma pesquisa de dois grupos com uma medida

O desenho desta pesquisa se caracteriza por duas amostras selecionadas aleatoriamente da mesma população para executarem dois experimentos X_1 e X_2 (um deles pode ser controle), cujo desfecho Y será avaliado em um único momento. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Tail(s); Effect size d; α err prob; Power (1 – β err prob)*** **e** ***Allocation ratio N2 / N1***. Preencha estes campos com os dados abaixo e vide a Figura 3:

Test Family = *t tests*

A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.

Statistical test = *Means: Difference between two independent means (two groups)*

O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.

Type of power analysis = *A priori: Compute required sample size – given α, power, and effect size*

O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.

Tail(s) = *Two ou One*

Two (duas caudas) devem ser selecionadas quando a diferença de médias entre os grupos pode ser tanto positiva como negativa. *One* (uma cauda) deve ser selecionada quando a diferença de médias entre os grupos pode ser somente positiva ou negativa.

Effect size d *= 0.20*

O tamanho do efeito pequeno é o recomendável para o cálculo amostral (vide Tabela 1).

α err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

Allocation ratio N2 / N1 *= 1*

É a relação entre as quantidades de participantes dos dois grupos determinada pela alocação aleatória para os grupos ou pelo pesquisador.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 788 participantes, sendo 394 para cada grupo, porque ***Allocation ratio N2 / N1*** = 1 (vide Figura 3).

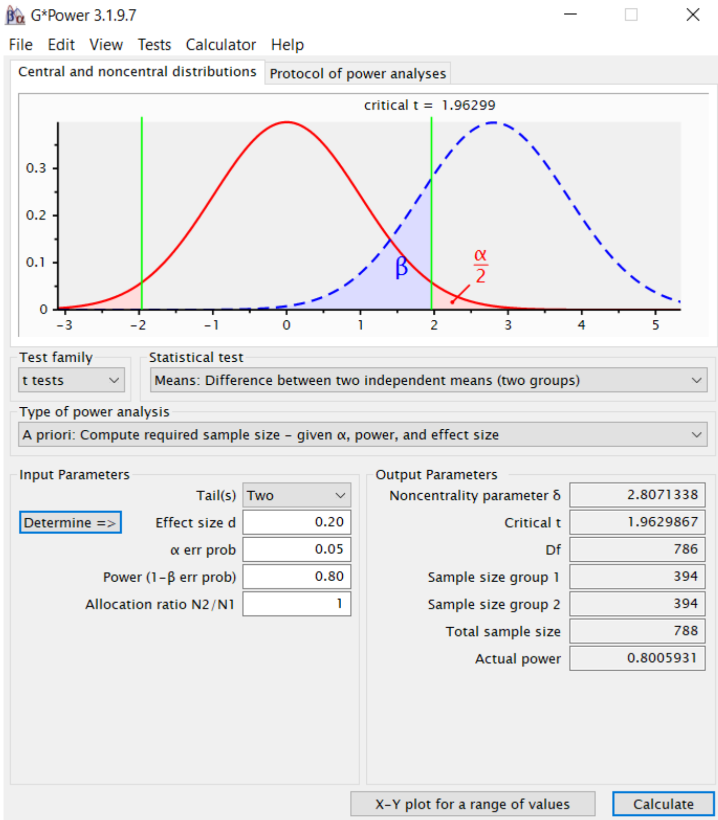


Figura 3 - Cálculo amostral para uma pesquisa de dois grupos com uma medida

Se você quer uma probabilidade pequena de estar errado (α = 0.05) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que o resultado da pesquisa não foi por acaso, ou seja, foi significativo, então você precisará de 788 participantes. Contudo, eu me solidarizo com a sua preocupação: será muito difícil recrutar 788 participantes para uma pesquisa. Mas há uma solução para isto: aumentando o tamanho do efeito d, terá como consequência um tamanho de amostra menor.

O tamanho do efeito d pode ser aumentado arbitrariamente ou determinado a partir dos resultados de um estudo piloto ou de uma pesquisa semelhante publicada. As duas últimas alternativas são muito melhores que a primeira. Supondo que os resultados de um estudo piloto tenham sido média do grupo 1 = 20; média do grupo 2 = 15; desvio padrão do grupo 1 = 7; desvio padrão do grupo 2 = 5.

Para determinar o tamanho do efeito d, clique em ***Determine*** = >, que abrirá um quadro secundário. Selecione ***n1 = n2,*** porque ***Allocation ratio N2 / N1*** *= 1*. Nos campos ***Mean group 1*** informe a média do grupo 1 = 20, ***Mean group 2*** informe a média do grupo 2 = 15; ***SD σ group 1*** informe o desvio padrão do grupo 1 = 7; ***SD σ group 2*** informe o desvio padrão do grupo 2 = 5; após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito d calculado no quadro secundário e igual a 0.8219949 foi transferido automaticamente para o campo ***Effect size d*** do quadro principal. Neste quadro, clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 50, sendo 25 para cada grupo (vide Figura 4). O G*Power calculou o tamanho do efeito d pela fórmula abaixo:

$$d = \frac{\text{Média do grupo 2} - \text{Média do grupo 1}}{\text{Raiz}\left(\dfrac{(\text{DP do grupo } 1^2 + \text{DP do grupo } 2^2)}{2}\right)}$$

sendo DP = desvio padrão.

O outro caminho para determinar o tamanho do efeito d no quadro secundário é ***n1 ≠ n2***, quando ***Allocation ratio N2 / N1*** ≠ *1*. Neste caminho, inserir no campo ***Mean group1*** a média do grupo 1, inserir no campo ***Mean group 2*** a média do grupo 2 e inserir no campo ***SD σ within each group*** o erro padrão da diferença. Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito d calculado no quadro secundário será transferido automaticamente para o campo ***Effect size d*** do quadro principal.

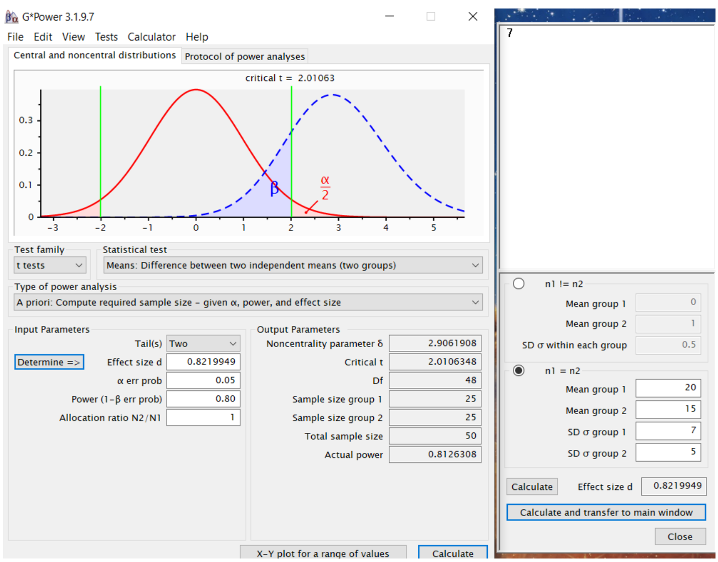


Figura 4 – Cálculo amostral para uma pesquisa de dois grupos com uma medida, determinando o tamanho do efeito d a partir dos resultados de um estudo piloto no quadro secundário

Tamanho da amostra para uma pesquisa de três ou mais grupos com uma medida

O desenho desta pesquisa se caracteriza por três ou mais grupos selecionados aleatoriamente da mesma população para executarem três ou mais experimentos X_1, X_2, X_3... X_n (um deles pode ser controle), cujo desfecho Y será avaliado em um único momento. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Effect size f; α err prob; Power (1 – β err prob)*** **e** ***Number of groups***. Preencha estes campos com os dados abaixo e vide a Figura 5:

Test Family *= F tests*

A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.

Statistical test *= ANOVA: Fixed effects, omnibus, one-way*

O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.

Type of power analysis *= A priori: Compute required sample size – given α, power, and effect size*

O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.

Effect size f *= 0.10*

O tamanho do efeito pequeno é o recomendável para o cálculo amostral (vide Tabela 1).

A err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

Number of groups *= 3*

É a quantidade de grupos existente no desenho da pesquisa.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 969 participantes (vide Figura 5).

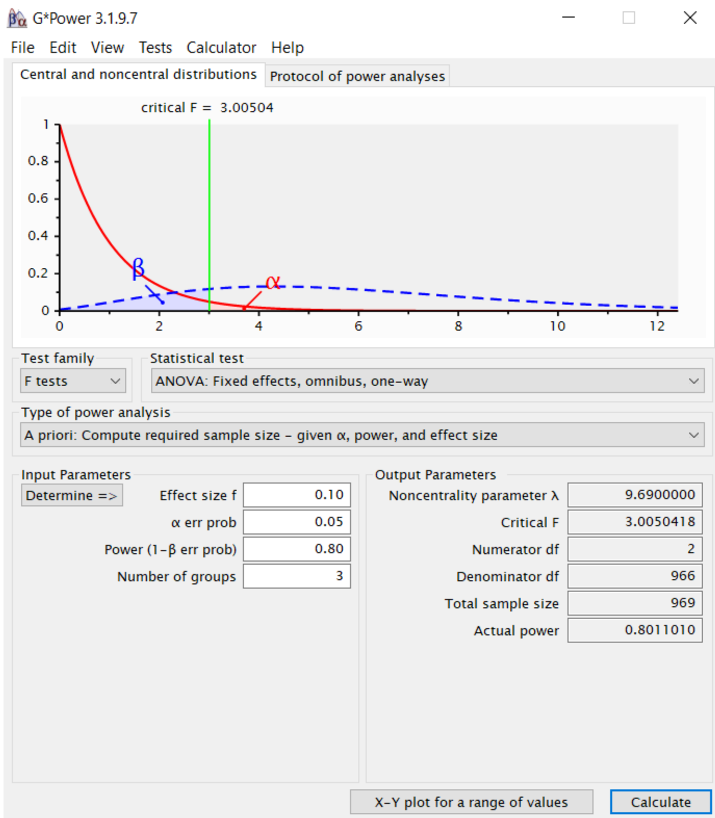


Figura 5 – Cálculo amostral para uma pesquisa de três grupos com uma medida

Se você quer uma probabilidade pequena de estar errado ($\alpha = 0.05$) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que o resultado da pesquisa não foi por acaso, ou seja, foi significativo, então você precisará de 969 participantes. Contudo, mais uma vez, eu me solidarizo com a sua preocupação: será muito difícil recrutar 969 participantes para uma pesquisa. Mas há uma solução para isto: aumentando o tamanho do efeito f, terá como consequência um tamanho de amostra menor.

O tamanho do efeito f pode ser aumentado arbitrariamente ou determinado a partir dos resultados de um estudo piloto ou de uma pesquisa semelhante publicada. As duas últimas alternativas são muito melhores que a primeira. Supondo que os resultados de uma pesquisa semelhante publicada tenham sido: média do grupo 1 = 20; média do grupo 2 = 15; média do grupo 3 = 10; desvio padrão do grupo 1 = 7; desvio padrão do grupo 2 = 5; desvio padrão do grupo 3 = 3.

Para determinar o tamanho do efeito f, clique em ***Determine*** = >, que abrirá um quadro secundário. Em ***Select procedure*** selecione ***Effect size from variance*** e ***From variances***. Nos campos ***Variance explained by special effect*** informe a variância dos efeitos, ***Variance within groups*** informe a variância dos erros.

A variância dos efeitos é obtida pela média dos quadrados dos desvios das médias dos grupos, em relação à média mãe, sendo a média mãe igual a média das médias dos grupos; a variância dos erros é obtida pela média dos quadrados dos desvios padrão dos grupos. Então,

Média mãe = (20 + 15 + 10) / 3 = 15

Variância dos efeitos = $(20 – 15)^2 + (15 – 15)^2 + (10 – 15)^2 / 3 = 16.7$

Variância dos erros = $(7^2 + 5^2 + 3^2) / 3 = 27.7$

Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito f calculado no quadro secundário e igual a 0.7764587 foi transferido automaticamente para o campo ***Effect size f*** do quadro principal. Neste quadro, clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 21 participantes (vide Figura 6). O G*Power calculou o tamanho do efeito f pela fórmula abaixo:

$$\text{Tamanho do efeito f} = \text{raiz}\left(\frac{\text{Variância dos efeitos}}{\text{Variância dos erros}}\right)$$

sendo a variância dos efeitos obtida pela média dos quadrados dos desvios das médias dos grupos, em relação à média mãe; sendo a média mãe igual a média das médias dos grupos; e a variância dos erros obtida pela média dos quadrados dos desvios padrão dos grupos.

O outro caminho para determinar o tamanho do efeito f no quadro secundário é ***Effect size from means***. Entretanto, considero este caminho não apropriado, porque depende de uma tabela de médias e tamanhos de grupos do G*Power 3.1.9.7 não fornecida e, principalmente, pela incongruência entre os dados informados no quadro principal e quadro secundário.

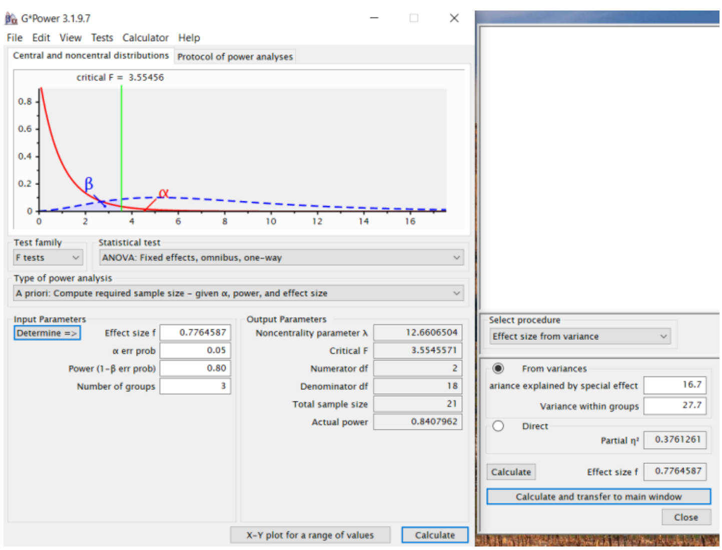


Figura 6 – Cálculo amostral para uma pesquisa de três grupos com uma medida, determinando o tamanho do efeito f a partir dos resultados de uma pesquisa semelhante no quadro secundário

Tamanho da amostra para uma pesquisa de um grupo com três ou mais medidas repetidas

O desenho desta pesquisa se caracteriza por um grupo selecionado aleatoriamente para executar um experimento X, cujo desfecho Y será avaliado em três ou mais momentos Y_1, Y_2, Y_3, ... Y_n. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Effect size f; α err prob; Power (1 – β err prob); Number of groups; Number of measurements; Corr among rep measures; Nonsphericity correction Ɛ*** e ***Option.*** Preencha estes campos com os dados abaixo e vide a Figura 7:

> ***Test Family*** *= F tests*
>
> A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.
>
> ***Statistical test*** *= ANOVA: Repeated measures, within factors*
>
> O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.
>
> ***Type of power analysis*** *= A priori: Compute required sample size – given α, power, and effect size*
>
> O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.
>
> ***Option*** = clique sobre este botão e selecione *as in G*Power 3.0*
>
> Apesar do G*Power recomendar a opção ***as in Cohen (1988)***, não considero esta uma boa opção, porque o

tamanho do efeito f será substituído pelo tamanho do efeito f(V), para o qual não há ainda classificação.

Effect size f *= 0.10*

O tamanho do efeito pequeno é o recomendável para o cálculo amostral (vide Tabela 1).

α err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a diferença entre as médias das medidas repetidas não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a diferença entre as médias das medidas repetidas não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

Number of groups *= 1*

É a quantidade de grupos existente no desenho da pesquisa.

Number of measurements *= 3*

É a quantidade de medidas repetidas do desfecho existente no desenho da pesquisa.

Corr among rep measures *= 0.5*

É a menor correlação da matriz r de Pearson entre as medidas repetidas.

Nonsphericity correction Ɛ *= 1*

É a correção da ausência de esfericidade entre as medidas repetidas, podendo variar de 1 / (***Number of measurements*** – 1) ≤ ***Ɛ*** ≤ 1. Quanto menor a correção, maior o tamanho amostral. No cálculo amostral, o valor 1 é o padrão utilizado.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 163 participantes (vide Figura 7).

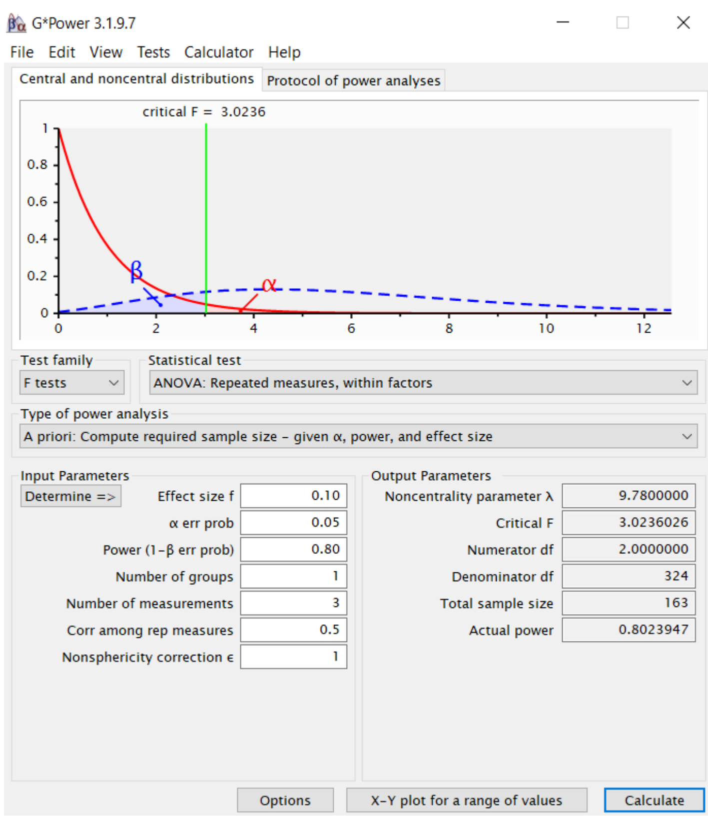


Figura 7 - Cálculo amostral para uma pesquisa de um grupo com três medidas repetidas

Se você quer uma probabilidade pequena de estar errado ($\alpha = 0.05$) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que o resultado da pesquisa não foi por acaso, ou seja, foi significativo, então você precisará de 163 participantes. Contudo, novamente, eu me solidarizo com a sua preocupação: será muito difícil recrutar 163 participantes para uma pesquisa. Mas há uma solução para isto: aumentando o tamanho do efeito f, terá como consequência um tamanho de amostra menor.

O tamanho do efeito f pode ser aumentado arbitrariamente ou determinado a partir dos resultados de um estudo piloto ou de uma pesquisa semelhante publicada. As duas últimas alternativas são muito melhores que a primeira. Supondo que os resultados de uma pesquisa semelhante publicada tenham sido: média do $Y_1 = 20$; média do $Y_2 = 15$; média do $Y_3 = 10$; desvio padrão do $Y_1 = 7$; desvio padrão do $Y_2 = 5$; desvio padrão do $Y_3 = 3$.

Para determinar o tamanho do efeito f, clique em ***Determine*** = >, que abrirá um quadro secundário. Selecione ***From variances***. Nos campos ***Variance explained by special effect*** informe a variância dos efeitos, ***Variance within groups*** informe a variância dos erros.

A variância dos efeitos é obtida pela média dos quadrados dos desvios das médias do desfecho nos diferentes momentos, em relação à média mãe, sendo a média mãe igual a média das médias do desfecho nos diferentes momentos; a variância dos erros é obtida pela média dos quadrados dos desvios padrão do desfecho nos diferentes momentos. Então,

Média mãe = $(20 + 15 + 10) / 3 = 15$

Variância dos efeitos = $(20 – 15)^2 + (15 – 15)^2 + (10 – 15)^2 / 3 = 16.7$

Variância dos erros = $(7^2 + 5^2 + 3^2) / 3 = 27.7$

Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito f calculado no quadro secundário e igual a 0.7764587 foi transferido automaticamente para o campo ***Effect size f*** do quadro principal. Neste quadro, clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 5 participantes (vide Figura 8). O G*Power calculou o tamanho do efeito f pela fórmula abaixo:

Variância dos efeitos
Tamanho do efeito f = raiz (--------------------------------------)
Variância dos erros

sendo a variância dos efeitos obtida pela média dos quadrados dos desvios das médias dos grupos, em relação à média mãe; sendo a média mãe igual a média das médias dos grupos; e a variância dos erros obtida pela média dos quadrados dos desvios padrão dos grupos.

O outro caminho para determinar o tamanho do efeito f no quadro secundário é ***Direct***. Neste caminho, inserir no campo ***Partial*** $\boldsymbol{\eta^2}$ = 0.01 (pequeno) ou 0.06 (médio) ou 0.14 (grande). Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito f calculado no quadro secundário será transferido automaticamente para o campo ***Effect size f*** do quadro principal.

Este outro caminho pelo ***Direct*** é o mesmo caminho que utiliza o tamanho do efeito f classificado na Tabela 1, porque η^2 = 0.01 (pequeno) corresponde a f = 0.10 (pequeno); η^2 = 0.06 (médio) corresponde a f = 0.25 (médio); η^2 = 0.14 (grande) corresponde a f = 0.40 (grande), já que $\eta^2 = f^2 / (1 + f^2)$.

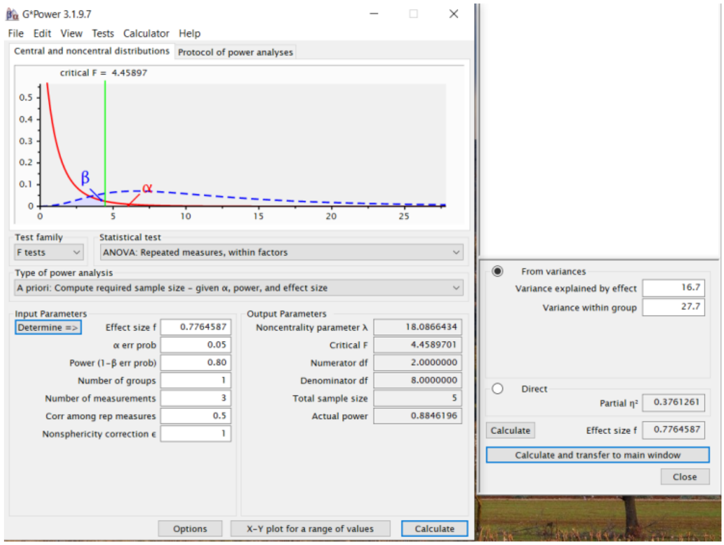


Figura 8 - Cálculo amostral para uma pesquisa de um grupo com três medidas repetidas, determinando o tamanho do efeito f a partir dos resultados de uma pesquisa semelhante no quadro secundário

Tamanho da amostra para uma pesquisa de dois ou mais grupos com duas ou mais medidas repetidas

O desenho desta pesquisa se caracteriza por dois ou mais grupos selecionados aleatoriamente da mesma população para executar um experimento X_1, X_2, ... X_n (um deles pode ser controle), cujo desfecho Y será avaliado em dois ou mais momentos Y_1, Y_2, ... Y_n. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Effect size f; α err prob; Power (1 – β err prob); Number of groups; Number of measruments; Corr among rep measures; Nonsphericity correction Ɛ*** e ***Option***. Preencha estes campos com os dados abaixo e vide a Figura 9:

Test Family *= F tests*

A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.

Statistical test *= ANOVA: Repeated measures, within-between interaction*

O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.

Type of power analysis *= A priori: Compute required sample size – given α, power, and effect size*

O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.

Option = clique sobre este botão e selecione *as in G*Power 3.0*

Apesar do G*Power recomendar a opção ***as in Cohen (1988)***, não a considero uma boa opção, porque o tamanho do efeito f será substituído pelo tamanho do efeito f(V), para o qual não há ainda classificação.

Effect size f *= 0.10*

O tamanho do efeito pequeno é o recomendável para o cálculo amostral (vide Tabela 1).

A err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

Number of groups *= 2*

É a quantidade de grupos existente no desenho da pesquisa.

Number of measurements *= 3*

É a quantidade de medidas repetidas do desfecho existente no desenho da pesquisa.

Corr among rep measures *= 0.5*

É a menor correlação da matriz r de Pearson entre as medidas repetidas.

Nonsphericity correction Ɛ *= 1*

É a correção da ausência de esfericidade entre as medidas repetidas, podendo variar de 1 / (***Number of measurements*** – 1) ≤ ***Ɛ*** ≤ 1. Quanto menor a correção, maior o tamanho amostral. No cálculo amostral, o valor 1 é o padrão utilizado.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 164 participantes (vide Figura 9).

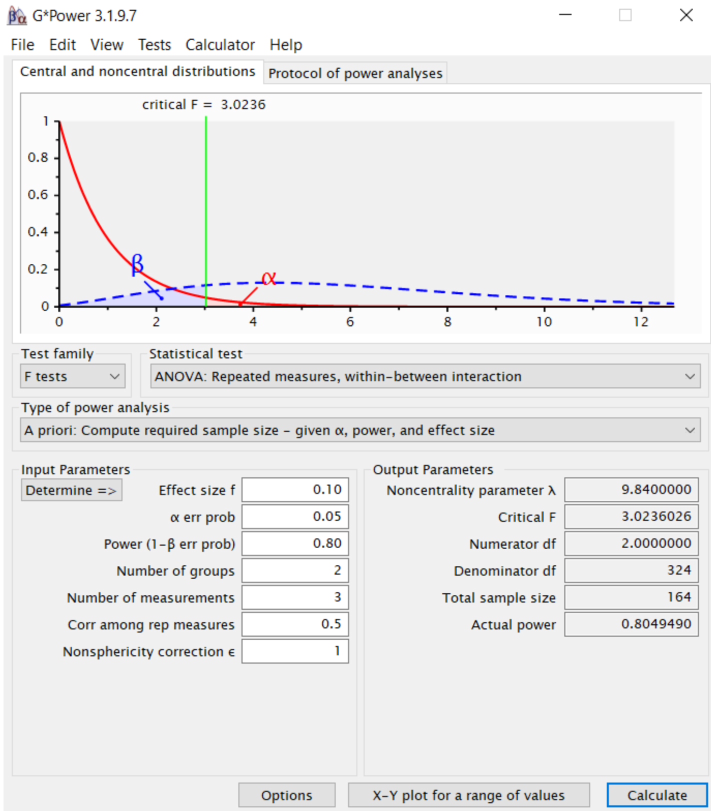


Figura 9 – Cálculo amostral para uma pesquisa de dois grupos com três medidas repetidas

Se você quer uma probabilidade pequena de estar errado (α = 0.05) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que o resultado da pesquisa não foi por acaso, ou seja, foi significativo, então você precisará de 164 participantes. Contudo, de novo, eu me solidarizo com a sua preocupação: será muito difícil recrutar 164 participantes para uma pesquisa. Mas há uma solução para isto: aumentando o tamanho do efeito f, terá como consequência um tamanho de amostra menor.

O tamanho do efeito f pode ser aumentado arbitrariamente ou determinado a partir dos resultados de um estudo piloto ou de uma pesquisa semelhante publicada. As duas últimas alternativas são muito melhores que a primeira. Supondo que os resultados em média ± desvio padrão de uma pesquisa semelhante publicada tenham sido:

Grupos	1º momento	2º momento	3º momento
1	20 ± 8	23 ± 10	25 ± 9
2	20 ± 9	19 ± 10	20 ± 11

x ± DP = média ± desvio padrão

Para determinar o tamanho do efeito f, clique em ***Determine*** = >, que abrirá um quadro secundário. Selecione ***From variances***. Nos campos ***Variance explained by special effect*** informe a variância dos efeitos, ***Variance within groups*** informe a variância dos erros.

A variância dos efeitos é obtida pela média dos quadrados dos desvios das médias do desfecho nos diferentes momentos, em relação à média mãe, sendo a média mãe igual a média das médias do desfecho nos diferentes momentos; a variância dos erros é obtida pela média dos quadrados dos desvios padrão do desfecho nos diferentes momentos. Então,

Média mãe = (20 + 23 + 25 + 20 + 19 + 20) / 6 = 21.2

Variância dos efeitos = $(20 - 21.2)^2 + (23 - 21.2)^2 + (25 - 21.2)^2 + (20 - 21.2)^2 + (19 - 21.2)^2 + (20 - 21.2)^2 / 6 = 4.5$

Variância dos erros = $(8^2 + 10^2 + 9^2 + 9^2 + 10^2 + 11^2) / 6 = 91.2$

Após a entrada destes dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito f calculado no quadro secundário e igual a 0.2221308 foi transferido automaticamente para o campo ***Effect size f*** do quadro principal. Neste quadro, clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 36 participantes (vide Figura 10). O G*Power calculou o tamanho do efeito f pela fórmula abaixo:

$$\text{Tamanho do efeito f} = \text{raiz}\left(\frac{\text{Variância dos efeitos}}{\text{Variância dos erros}}\right)$$

sendo a variância dos efeitos obtida pela média dos quadrados dos desvios das médias dos grupos, em relação à média mãe; sendo a média mãe igual a média das médias dos grupos; e a variância dos erros obtida pela média dos quadrados dos desvios padrão dos grupos.

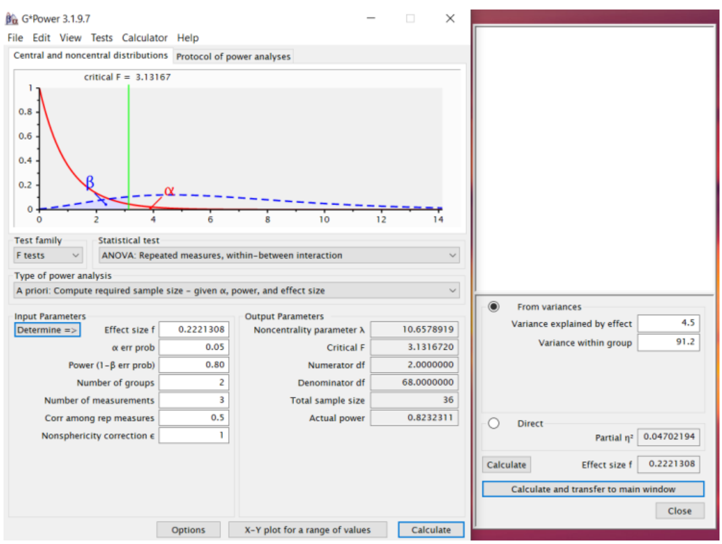


Figura 10 – Cálculo amostral para uma pesquisa de dois grupos com três medidas repetidas, determinando o tamanho do efeito f a partir dos resultados de uma pesquisa semelhante no quadro secundário

O outro caminho para determinar o tamanho do efeito f no quadro secundário é ***Direct***. Neste caminho, inserir no campo ***Partial*** $\boldsymbol{\eta^2}$ = 0.01 (pequeno) ou 0.06 (médio) ou 0.14 (grande). Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito f calculado no quadro secundário será transferido automaticamente para o campo ***Effect size f*** do quadro principal.

Este outro caminho pelo ***Direct*** é o mesmo caminho que utiliza o tamanho do efeito f classificado na Tabela 1, porque $\eta^2 = 0.01$ (pequeno) corresponde a f = 0.10 (pequeno); $\eta^2 = 0.06$ (médio) corresponde a f = 0.25 (médio); $\eta^2 = 0.14$ (grande) corresponde a f = 0.40 (grande), já que $\eta^2 = f^2 / (1 + f^2)$.

Cálculo amostral para testar a relação entre duas ou mais variáveis

Diferentes desenhos de pesquisas podem surgir da relação entre duas ou mais variáveis, que dependerá do objetivo da pesquisa de correlacionar duas variáveis ou predizer uma delas a partir das outras. No caso de predizer uma delas a partir das outras, se deve ainda considerar a natureza da variável a ser predita: contínua/intervalar, nominal que expressa uma dicotomia ou nominal que expressa uma quantidade. Os desenhos de pesquisas possíveis são: (1) correlação entre duas variáveis; (2) predição de uma variável contínua intervalar a partir de duas ou mais variáveis; (3) de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis; (4) de predição de uma variável nominal, que expressa uma quantidade, a partir de duas ou mais variáveis.

Tamanho da amostra para uma pesquisa de correlação entre duas variáveis

O desenho desta pesquisa se caracteriza por uma amostra selecionada aleatoriamente da qual se mede duas variáveis Y e W contínuos intervalares com distribuição normal. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Correlation ρ H_1; α err prob; Power (1 – β err prob); Correlation ρ H_0*** e ***Option****.* Preencha estes campos com os dados abaixo e vide a Figura 11:

Test Family *= Exact*

A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.

Statistical test *= Correlation: Bivariate normal model*

O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.

Type of power analysis *= A priori: Compute required sample size – given α, power, and effect size*

O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.

Option = clique sobre este botão e selecione *use* ***exact distribution if N < 10000*** ou ***use large sample approximation (Fisher Z)***

Para ambas as opções, os dados de entrada e o tamanho da amostra serão os mesmos.

Tail(s) *= Two ou One*

Two (duas caudas) devem ser selecionadas quando a correlação entre as variáveis pode ser tanto positiva como negativa. *One* (uma cauda) deve ser selecionada quando a correlação pode ser somente positiva ou negativa.

Correlation ρ H_1 *= 0.10*

É a menor correlação r de Pearson esperada para a hipótese de pesquisa (H_1) (vide Tabela 1).

α err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a correlação entre as duas variáveis não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β err prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a correlação entre as duas variáveis não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

Correlation ρ $\boldsymbol{H_0}$ *= 0*

É a correlação r de Pearson esperada para a hipótese nula, normalmente igual a zero.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 782 participantes (vide Figura 11).

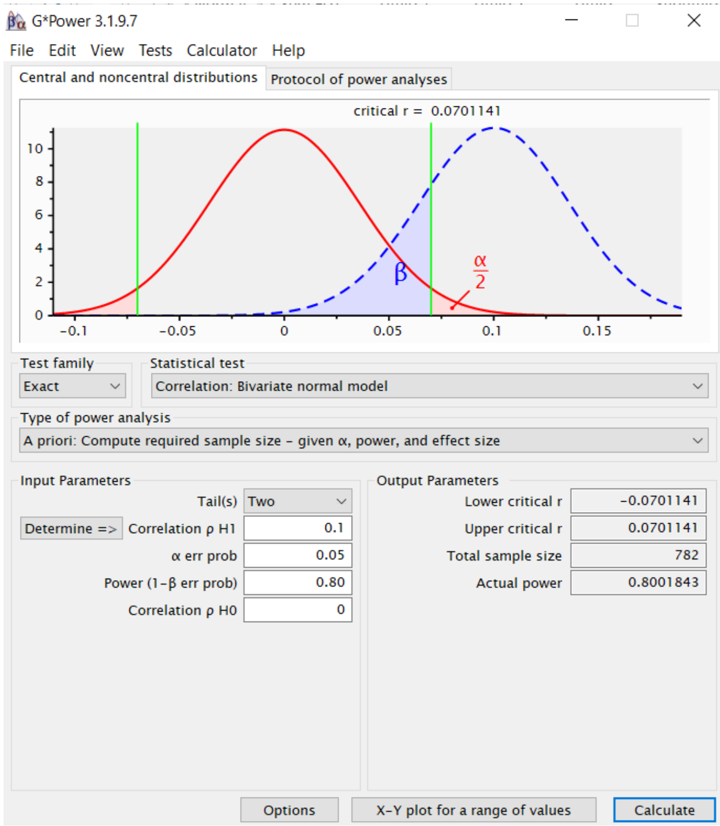


Figura 11 – Cálculo amostral para uma pesquisa de correlação entre duas variáveis

Se você quer uma probabilidade pequena de estar errado ($\alpha = 0.05$) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que o resultado da pesquisa não foi por acaso, ou seja, foi significativo, então você precisará de 782 participantes. Contudo, repetidamente, eu me solidarizo com a sua preocupação: será muito difícil recrutar 782 participantes para uma pesquisa. Mas há uma solução para isto: aumentando a correlação ρ H_1 entre as variáveis, terá como consequência um tamanho de amostra menor.

A correlação ρ H_1 pode ser aumentada arbitrariamente ou determinada a partir dos resultados de um estudo piloto. A segunda alternativa é mais interessante, é mais científica, que a primeira. Evidentemente, que a quantidade de participantes do estudo piloto será pequena. Por isto mesmo, ignore o valor-P encontrado. Certamente, o valor-P será maior que 0.05, porque a amostra será muito pequena. Então, considere somente o r de Pearson obtido. Muito provavelmente, o r de Pearson obtido no estudo piloto será maior que a correlação ρ $H_1 = 0.10$ utilizado para o cálculo amostral de 782 participantes. Supondo que o r de Pearson do estudo piloto tenha sido 0.6, o tamanho da amostra será de 19 participantes (vide Figura 12).

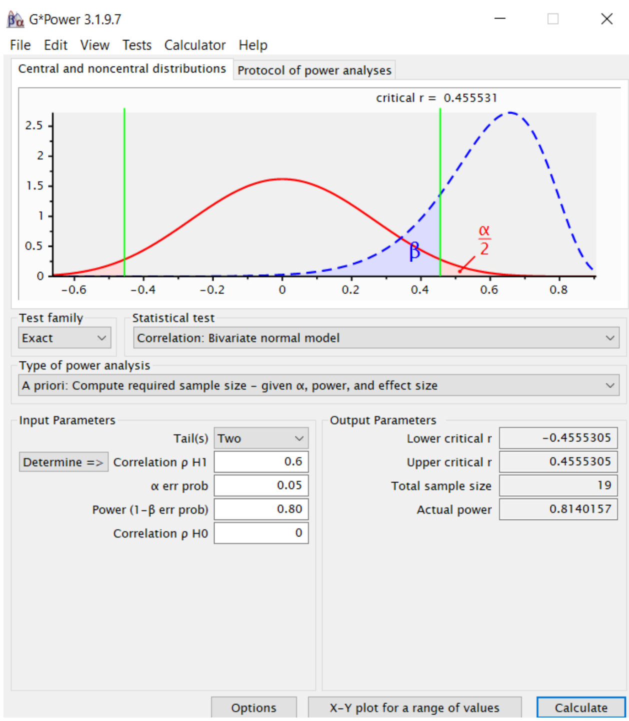


Figura 12 – Cálculo amostral para uma pesquisa de correlação entre duas variáveis, utilizando o r de Pearson de um estudo piloto no quadro secundário

Para este desenho da pesquisa, associação entre duas variáveis contínuas, a ***Correlation ρ H₁***, ou seja, o coeficiente de correlação r de Pearson, assemelha-se ao tamanho do efeito (***Effect size***) presente nos desenhos de pesquisa de diferença de médias: quanto maior a ***Correlation ρ H₁***, menor será o tamanho da amostra.

Tamanho da amostra para uma pesquisa de predição de uma variável contínua intervalar a partir de duas ou mais variáveis

O desenho desta pesquisa se caracteriza por uma amostra selecionada aleatoriamente da qual se mede a variável contínua intervalar predita Y e duas ou mais variáveis contínuas intervalares ou nominais preditoras X_1, X_2 ... X_n. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Effect size f²; α err prob; Power (1 – β err prob);*** e ***Number of predictors****.* Preencha estes campos com os dados abaixo e vide a Figura 13:

> ***Test Family*** *= F tests*
>
> A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.
>
> ***Statistical test*** *= Linear multiple regression: Fixed model,* R^2 *deviation from zero*
>
> O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.

***Type of power analysis** = A priori: Compute required sample size – given α, power, and effect size*

O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.

***Effect size** f^2 = 0.02*

O tamanho do efeito pequeno é o recomendável para o cálculo amostral (vide Tabela 1).

***α err prob** = 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a predição da variável Y não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

***Power (1 – β err prob)** = 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a predição da variável Y não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

***Number of predictor** = 2*

É a quantidade de variáveis preditoras previstas no desenho da pesquisa.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 485 participantes (vide Figura 13).

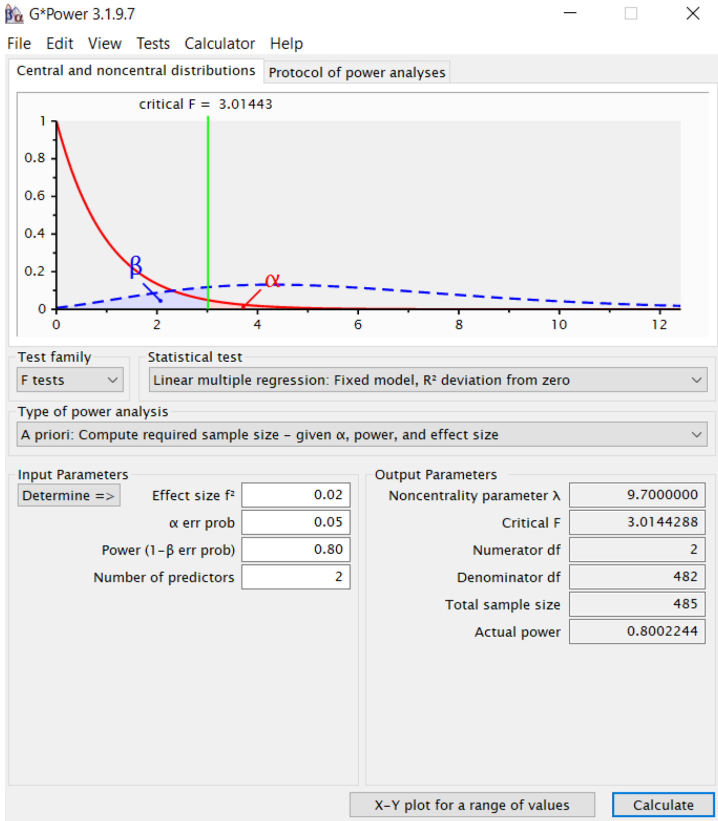


Figura 13 - Cálculo amostral para uma pesquisa de predição de uma variável contínua intervalar a partir de duas variáveis

Se você quer uma probabilidade pequena de estar errado ($\alpha = 0.05$) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que o resultado da pesquisa não foi por acaso, ou seja, foi significativo, então você precisará de 485 participantes. Contudo, novamente, eu me solidarizo com a sua preocupação: será muito difícil recrutar 485 participantes para uma pesquisa. Mas há uma solução para isto: aumentando o tamanho do efeito f^2, terá como consequência um tamanho de amostra menor.

O tamanho do efeito f^2 pode ser aumentado arbitrariamente ou determinado a partir dos resultados de um estudo piloto. A segunda alternativa é mais interessante, é mais científica, que a primeira. Evidentemente, que a quantidade de participantes do estudo piloto será pequena: recomendo 5 participantes por variável preditora. Por isto mesmo, ignore o valor-P encontrado. Certamente, o valor-P será maior que 0.05, porque a amostra será muito pequena. Então, considere somente o coeficiente de determinação R^2 obtido. Muito provavelmente, o coeficiente de determinação R^2 obtido no estudo piloto será maior que aquele correspondente ao tamanho do efeito f^2 igual a 0.02 que redundou na amostra de 485 participantes. Supondo que o R^2 do estudo piloto tenha sido 0.23, clique em ***Determine*** = >, que abrirá um quadro secundário. Selecione ***From correlation coefficient.*** **No campo *Squared multiple correlation ρ²***, informe o coeficiente de determinação R^2 do estudo piloto de 0.23. Após a entrada deste dado, clique em ***Calculate and transfer to main window***. O tamanho do efeito f^2 calculado no quadro secundário e igual a 0.2987013 foi transferido automaticamente para o campo ***Effect size f²*** do quadro principal. Neste quadro, clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 36 participantes (vide Figura 14). O G*Power calculou o tamanho do efeito f^2 pela fórmula abaixo:

$$\text{Tamanho do efeito } f^2 = \frac{R^2}{1 - R^2}$$

sendo R^2 igual ao coeficiente de determinação.

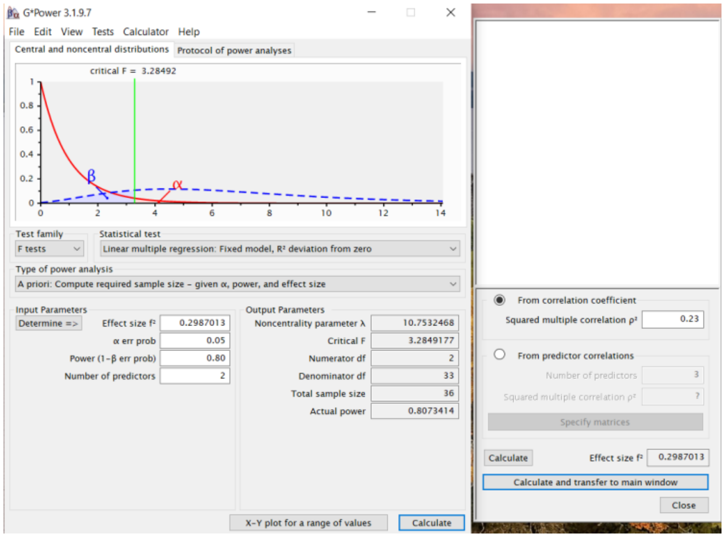


Figura 14 - Cálculo amostral para uma pesquisa de predição de uma variável contínua intervalar a partir de duas variáveis, utilizando o coeficiente de determinação R^2 de um estudo piloto no quadro secundário

O outro caminho para determinar o tamanho do efeito f^2 no quadro secundário é ***From predictor correlations***. Neste caminho, clicar sobre ***Specify matrices***. Na aba ***Corr between predictors and outcome*** inserir as correlações r de Pearson entre as variáveis preditoras e a variável predita. Na aba ***Corr between predictors*** inserir as correlações de Pearson entre as variáveis preditoras. Em seguida, clicar em ***Accept values***. Após a entrada dos dados, clique em ***Calculate and transfer to main window***. O tamanho do efeito f^2 calculado no quadro secundário será transferido automaticamente para o campo ***Effect size f^2*** do quadro principal.

Tamanho da amostra para uma pesquisa de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis

O desenho desta pesquisa se caracteriza por uma amostra selecionada aleatoriamente da qual se mede a variável nominal dicotômica predita Y, que expressa a ocorrência ou a não ocorrência de um evento, e duas ou mais variáveis contínuas intervalares ou nominais preditoras X_1, X_2 ... X_n com distribuição normal ou discreta. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Tail(s); Odds ratio; Pr (Y = 1) (X = 1) H_0; α err prob; Power (1 – β err prob); R^2 other X; X distribution; X parm μ; X parm σ*** e ***Option****.* Preencha estes campos com os dados abaixo e vide a Figura 15:

> ***Test Family*** *= z tests*
>
> A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.
>
> ***Statistical test*** *= Logistic regression*
>
> O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.
>
> ***Type of power analysis*** *= A priori: Compute required sample size – given α, power, and effect size*
>
> O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.
>
> ***Option*** = clique sobre este botão e selecione ***as Input effect size as Odds ratio*** e ***Use large sample approximation Test***

procedure Demidenko (2007) - recommended e ***With variance correction.***

É o recomendável.

Tail(s) *= Two ou One*

Two (duas caudas) devem ser selecionadas quando a predição da variável Y pode tanto ocorrer como não ocorrer. *One* (uma cauda) deve ser selecionada quando a predição da variável Y pode somente ocorrer ou não ocorrer.

Odds ratio *= 2*

A ***Odds ratio*** representa o tamanho do efeito. Por isto mesmo, a odds ratio pequena é a recomendável para o cálculo amostral. O *Grading of Recommendations Assessment, Development and Evaluation* (GRADE) classifica como pequena uma odds ratio ≤ 2. Os limites são de 0 ≤ ***Odds ratio*** ≤ 100.

Pr (Y = 1) (X = 1) $\boldsymbol{H_0}$ *= 0.5*

Probabilidade de ocorrência da hipótese nula (H_0), ou seja, do evento não ocorrer, podendo variar de 0 ≤ ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_0}$ ≤ 1. Quando 0.1 ≤ ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_0}$ ≤ 0.5 ou 0.5 ≤ ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_0}$ ≤ 0.9, o tamanho da amostra tende a aumentar. Por isto, 0.5 é o recomendável.

α err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a predição da variável Y não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β err prob) *= 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a predição da variável Y não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

R²* *other X *= 0.40*

É um fator de correção da influência das outras variáveis preditoras, no poder do teste (1 – erro β), para modelos com mais de uma variável preditora. Na prática, o ***R² other X*** é igual ao R^2 do modelo que prediz a variável preditora com o menor coeficiente de regressão, tendo como variáveis preditoras as outras variáveis preditoras do modelo. Quanto maior o ***R² other X*** maior será o tamanho da amostra. Se n é o tamanho da amostra considerando a variável preditora com o menor coeficiente de regressão sozinha, então o tamanho da amostra em um modelo com variáveis preditoras adicionais é: $n' = n / (1 - R^2)$. 0 ≤ ***R² other X*** < 1 e assume o valor 0, quando tiver somente uma variável preditora. A meu ver, ***R² other X*** expressa a multicolinearidade entre as variáveis preditoras do modelo, que deve ser evitada em uma análise de regressão.

X distribution *= Normal*

É o tipo de distribuição das variáveis preditoras ou independentes ou covariáveis. Se as variáveis preditoras ou independentes ou covariáveis forem discretas, então o tipo de distribuição poderá ser binomial (dicotômica) ou Poisson (quantidade). Se as variáveis preditoras ou independentes ou covariáveis forem contínuas intervalares, então o tipo de distribuição poderá ser normal, lognormal, exponencial ou uniforme (dependerá da característica contínua destas variáveis).

X parm μ = *0*

A média de uma distribuição normal é zero.

X parm σ = *1*

O desvio padrão de uma distribuição normal é um, na qual 68% dos resultados estão distribuídos entre ± 1 desvio padrão.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 136 participantes (vide Figura 15).

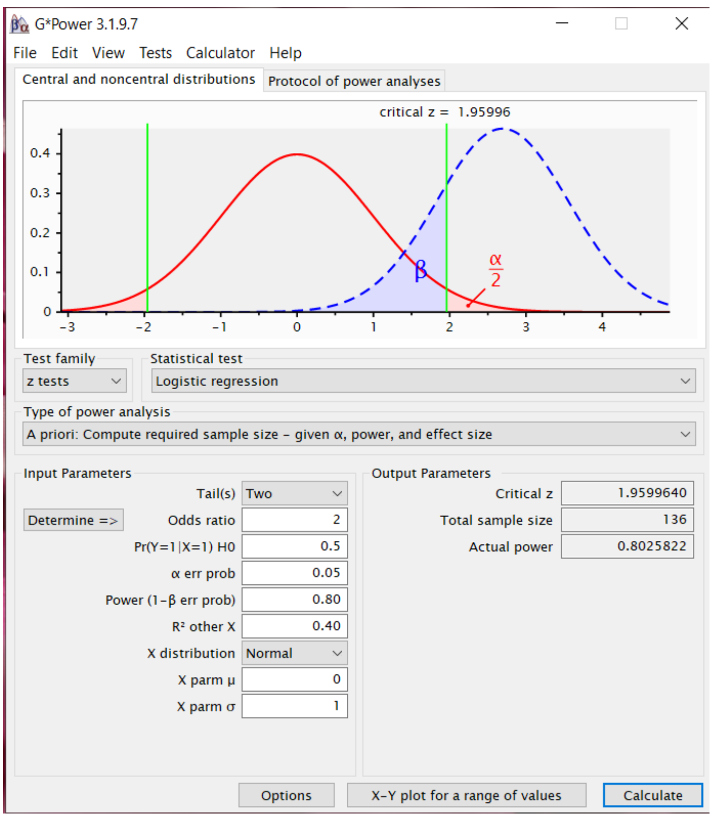


Figura 15 - Cálculo amostral para uma pesquisa de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis

Se você quer uma probabilidade pequena de estar errado ($\alpha = 0.05$) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que a predição da variável Y não foi por acaso, ou seja, foi significativa, então você precisará de 136 participantes. Contudo, de novo, eu me solidarizo com a sua preocupação: será muito difícil recrutar 136 participantes para uma pesquisa. Mas há uma solução para isto: refinar a ***Odds ratio*** e o ***R^2 other X***.

A ***Odds ratio*** e o ***R^2 other X*** podem ser refinados a partir dos resultados de um estudo piloto. A quantidade de participantes do estudo piloto será pequena: recomendo 5 participantes por variável X preditora. Por isto mesmo, ignore o valor-P encontrado. Certamente, o valor-P será maior que 0.05, porque a amostra será muito pequena. Supondo que o estudo piloto, que espelha seu estudo principal, contenha três variáveis preditoras e o menor coeficiente de regressão tenha sido de 4.5, o R^2 do modelo de predição da variável preditora com o menor coeficiente de regressão, tendo como variáveis preditoras as outras duas variáveis preditoras, tenha sido 0.30. Então, o ***R^2 other X*** será igual a 0.30 e a ***Odds ratio*** será igual 4.5. O tamanho da amostra será de 42 participantes (vide Figura 16).

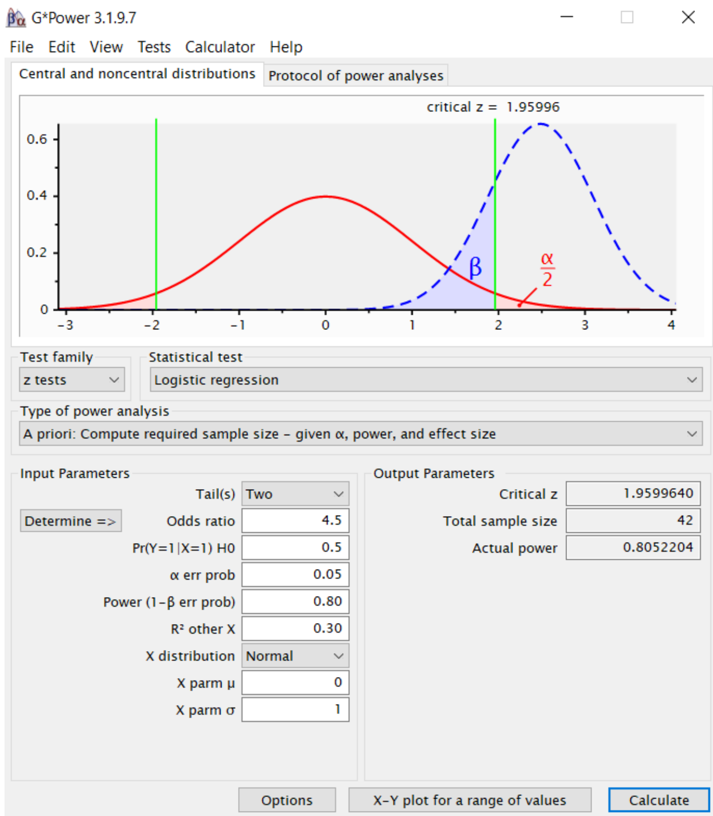


Figura 16 - Cálculo amostral para uma pesquisa de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis, utilizando a *Odds ratio* e o R^2 *other X* de um estudo piloto

Há um outro caminho para o cálculo da amostra para uma pesquisa de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis preditoras. Ele é semelhante ao visto acima. A única diferença é que o tamanho do efeito expresso pela ***Odds ratio*** será determinado pelas probabilidades de o evento ocorrer ***Pr (Y = 1) (X = 1)*** H_1 e de não ocorrer ***Pr (Y = 1) (X = 1)*** H_0.

Para isto, clique em ***Determine*** =>. Abrirá o quadro secundário para que sejam informadas as probabilidades do evento ocorrer ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_1}$ e não ocorrer ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_0}$. Para o exemplo dado na Figura 16, a probabilidade de o evento ocorrer ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_1}$ = 0.818182 e de não ocorrer ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_0}$ = 0.5. Clique em ***Calculate and tranfer to main window***. Retorne ao quadro principal e clique em ***Calculate*** (vide Figura 17).

O campo ***Odds ratio*** no quadro secundário não pode ser preenchido. Ele é calculado pela razão das chances entre as duas probabilidades ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_1}$ / ***Pr (Y = 1) (X = 1)*** $\boldsymbol{H_0}$. A chance de o evento ocorrer foi de 0.818182 / (1 – 0.818182) = 4.5. A chance de o evento não ocorrer foi de 0.5 / (1 – 0.5) = 1. A ***Odds ratio*** foi igual a 4.5 / 1 = 4.5.

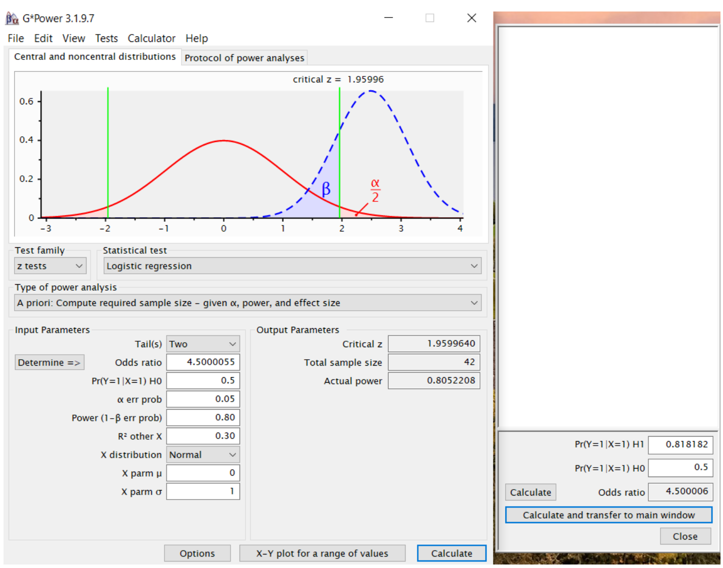


Figura 17 - Cálculo amostral para uma pesquisa de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis, utilizando as probabilidades de o evento ocorrer e de o evento não ocorrer

Tamanho da amostra para uma pesquisa de predição de uma variável nominal, que expressa uma quantidade, a partir de duas ou mais variáveis

O desenho desta pesquisa se caracteriza por uma amostra selecionada aleatoriamente da qual se mede a variável nominal predita Y, que expressa uma quantidade, e duas ou mais variáveis preditoras contínuas intervalares ou nominais X_1, X_2 ... X_n. Qual deveria ser o tamanho desta amostra?

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Tail(s); Exp (β₁); α err prob; Power (1 – β err prob); Base rate exp (β₀); Mean exposure; R² other X; X distribution; X parm μ; X parm σ*** e ***Option***. Preencha estes campos com os dados abaixo e vide a Figura 18:

> ***Test Family*** *= z tests*
>
> A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.
>
> ***Statistical test*** *= Poisson regression*
>
> O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.
>
> ***Type of power analysis*** *= A priori: Compute required sample size – given α, power, and effect size*
>
> O tamanho da amostra estimado depende do erro α, do poder (1 – erro β) e do tamanho do efeito informados.
>
> ***Option*** = clique sobre este botão e selecione ***Use large sample approximation Test procedure Demidenko (2007)*** e ***With variance correction.***

É recomendável.

Tail(s) *= Two ou One*

Two (duas caudas) devem ser selecionadas quando a predição da variável Y pode estar tanto acima como abaixo da taxa média de eventos. *One* (uma cauda) deve ser selecionada quando a predição da variável Y pode somente estar acima ou abaixo da taxa média de eventos.

Exp (β_1) *= 1,1*

É o aumento relativo da taxa de eventos sobre a taxa média de eventos, ***Base rate exp (β_0),*** assumida em hipótese nula (H_0), se a variável preditora for aumentada de uma unidade. Por exemplo, se um aumento de 10% sobre a taxa média de eventos for assumido, quando a variável preditora for aumentada em uma unidade, ***Exp (β_1)*** será definido como 1,1. Uma entrada de ***Exp (β_1)*** = 1 corresponde a "sem efeito" e não deve ser usado em cálculos a priori. Seus valores variam de 0 < ***Exp (β_1)*** ≤ 1000.

α err prob *= 0.05*

Caso a amostra estimada seja alcançada na pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a predição da variável Y não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado no cálculo amostral.

Power (1 – β err prob*) = 0.80*

Caso a amostra estimada seja alcançada na pesquisa, 0.80 será o poder (1 – erro β) mínimo obtido para o pesquisador afirmar que a predição da variável Y não foi devido ao acaso. Poderes maiores que este vão requerer amostras maiores. Por isto, 0.80 é o padrão a ser utilizado no cálculo amostral.

Base rate exp* (β_0) *= 5

Esta é a taxa média de eventos assumida em hipótese nula (H_0), que significa a taxa de eventos ocorrida no grupo não exposto, ou ainda a taxa média de eventos ocorridos em um determinado tempo, período ou situação. Deve ser maior que 0.

Mean exposure *= 1*

Esta é a unidade de tempo durante a qual os eventos são contados. Deve ser maior que 0.

R^2 other X *= 0.3*

É um fator de correção da influência das outras variáveis preditoras, no poder do teste (1 – erro β), para modelos com mais de uma variável preditora. Na prática, o ***R^2 other X*** é igual ao R^2 do modelo que prediz a variável preditora com o menor coeficiente de regressão, tendo como variáveis preditoras as outras variáveis preditoras. Quanto maior o ***R^2 other X*** maior será o tamanho da amostra. Se n é o tamanho da amostra considerando a variável preditora com o menor coeficiente de regressão sozinha, então o tamanho da amostra em um modelo com variáveis preditoras adicionais é: $n' = n / (1 - R^2)$. $0 \leq$ ***R^2 other X*** < 1 e assume o valor 0, quando tiver somente uma variável preditora. A meu ver, ***R^2 other X*** expressa a multicolinearidade entre as variáveis preditoras do modelo, que deve ser evitada em uma análise de regressão.

X distribution *= Binomial*

É o tipo de distribuição das variáveis preditoras ou independentes ou covariáveis. Se as variáveis preditoras ou independentes ou covariáveis forem discretas, então o tipo de distribuição poderá ser binomial (dicotômica) ou Poisson (quantidade). Se as variáveis preditoras ou independentes ou covariáveis forem contínuas intervalares, então o tipo de distribuição poderá ser normal, lognormal, exponencial ou uniforme (dependerá da característica contínua destas variáveis).

X parm π *= 0.5*

A média de uma distribuição binomial é 0.5.

Clique em ***Calculate*** e veja o tamanho da amostra no campo ***Total sample size*** = 942 participantes.

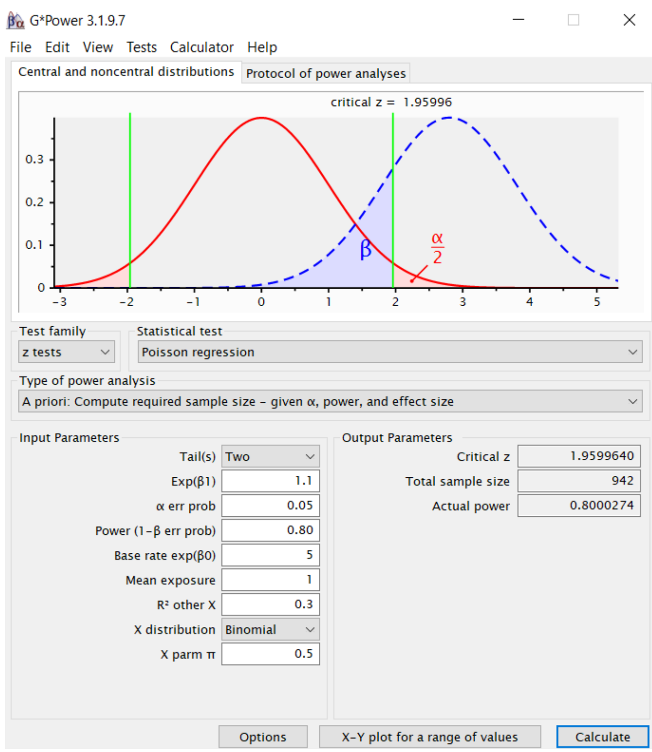


Figura 18 - Cálculo amostral para uma pesquisa de predição de uma variável nominal, que expressa uma quantidade, que expressa uma quantidade, a partir de duas ou mais variáveis

Se você quer uma probabilidade pequena de estar errado ($\alpha = 0.05$) e uma probabilidade grande de estar certo (poder = 0.80) ao afirmar que a predição da variável Y não foi por acaso, ou seja, foi significativa, então você precisará de 942 participantes. Contudo, de novo, eu me solidarizo com a sua preocupação: será muito difícil recrutar 942 participantes para uma pesquisa. Mas há uma solução para isto: refinar o ***Exp (β_1)*** e o ***R^2 other X***.

O ***Exp (β_1)*** e o ***R^2 other X*** podem ser refinados a partir dos resultados de um estudo piloto. A quantidade de participantes do estudo piloto será pequena: recomendo 5 participantes por variável X preditora. Por isto mesmo, ignore o valor-P encontrado. Certamente, o valor-P será maior que 0.05, porque a amostra será muito pequena. Supondo que o estudo piloto, que espelha seu estudo principal, contenha três variáveis preditoras, ocorreu um aumento de 40% sobre a taxa média de eventos, quando a variável preditora foi aumentada em uma unidade e o R^2 do modelo de predição da variável com o menor coeficiente de regressão, tendo como variáveis preditoras as outras duas variáveis preditoras, tenha sido 0.1. Então, o ***R^2 other X*** será igual a 0.1 e a ***Exp (β_1)*** será igual 1.4. O tamanho da amostra será de 53 participantes (vide Figura 19).

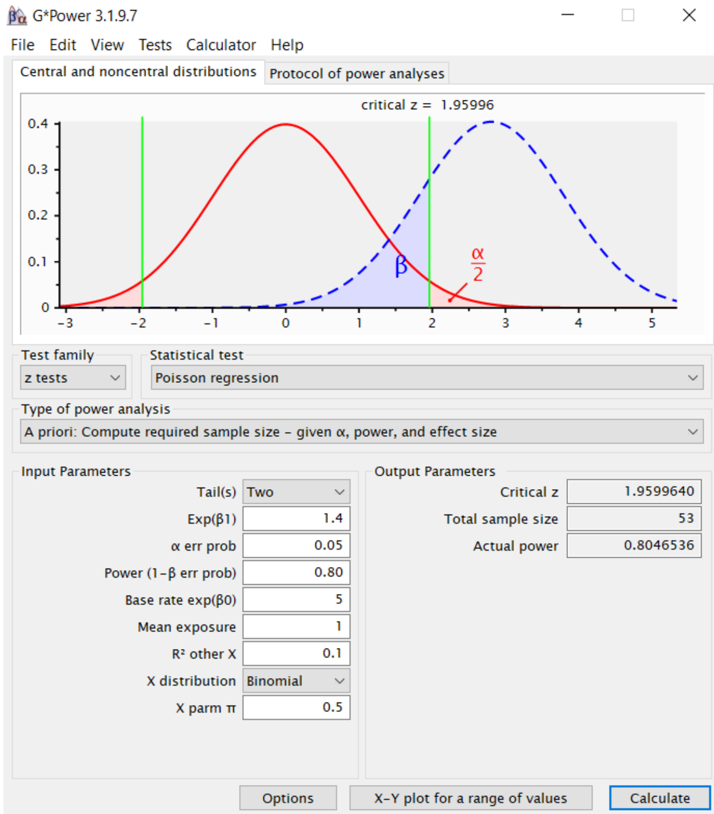


Figura 19 - Cálculo amostral para uma pesquisa de predição de uma variável nominal, que expressa uma quantidade, a partir de duas ou mais variáveis, utilizando o Exp (β1) e o R^2 *other X* de um estudo piloto

Nos oito desenhos de pesquisas da Tabela 2, o tamanho da amostra inicial considerou os tamanhos dos efeitos "pequenos" TE_1 (Tabela 1), que redundou em tamanhos de amostras grandes n_1. Mantidos todos os dados dos campos do G*Power 3.1.4.1 deste cálculo amostral e substituindo somente o tamanho de efeito TE_1 pelos tamanhos dos efeitos obtidos no estudo piloto ou estudo semelhante publicado TE_2, o tamanho da amostra diminuiu consideravelmente para n_2. Isto mostra quanto o tamanho do efeito interfere no tamanho da amostra: quanto o maior o tamanho do efeito, menor será o tamanho amostral.

Tabela 2 – Sumário comparativo da diminuição do tamanho da amostra de n_1 para n_2, devido ao aumento do tamanho do efeito de TE_1 para TE_2 ocorrido nos diversos desenhos das pesquisas

Desenhos das pesquisas	**n_1**	**n_2**	**TE_1**	**TE_2**
2 grupos e uma medida (Fig. 3 e 4)	788	50	d =0.20	d = 0.82
3^+ grupos e uma medida (Fig. 5 e 6)	969	21	f = 0.10	f = 0.77
1 grupo e três^+ medidas repetidas (Fig. 7 e 8)	163	5	f = 0.10	f = 0.77
2^+ grupos e duas^+ medidas repetidas (Fig. 9 e 10)	164	36	f = 0.10	f = 0.22
Correlação de duas variáveis (Fig. 11 e 12)	782	19	ρ = 0.10	ρ = 0.60
Predição de uma variável contínua intervalar (Fig. 13 e 14)	485	36	f^2 = 0.02	f^2 = 0.29
Predição de uma variável dicotômica (Fig. 15 e 16)	136	42	OR = 2 R^2 other X = 0.4	OR = 4.5 R^2 other X = 0.3
Predição de uma variável de quantidade (Fig. 18 e 19)	942	53	$Exp(\beta_1)$ = 1.1 R^2 other X = 0.3	$Exp(\beta_1)$ = 1.4 R^2 other X = 0.1

Nem sempre o tamanho da amostra estimado é alcançado ao final da pesquisa: o que fazer?

Algumas vezes, o tamanho da amostra estimado no cálculo amostral pode não ser obtido no final da pesquisa. Isto se deve a diversas causas: (1) a dificuldade intrínseca para recrutar pessoas interessadas em participar da pesquisa; (2) a mortalidade experimental que ocorre durante a pesquisa, porque os participantes desistem de continuar devido ao surgimento de efeitos colaterais, ou conflitos de interesses com outras atividades que surgiram, ou desmotivação; (3) a um período de tempo de coleta de dados pré-determinado, por exemplo, nos cursos de mestrado este tempo é de aproximadamente 6 a 8 meses, enquanto nos cursos de doutorado de 1 ano a 1 ano e 4 meses. O que fazer?

Uma solução interessante para quando ocorrer esta situação é analisar os dados considerando a hipótese nula (H_0) falsa, diferentemente do contumaz que é analisar os dados considerando a hipótese nula (H_0) verdadeira. Normalmente, a análise dos dados considerando a hipótese nula (H_0) verdadeira se baseia no erro α, erro máximo aceitável para o pesquisador afirmar que o resultado obtido na pesquisa não foi devido ao acaso, menor ou igual a 5%. Ao invés disto, a análise dos dados considerando a hipótese nula falsa se baseia no poder (1 – erro β), o acerto mínimo que tem o pesquisador para afirmar que o resultado da pesquisa não foi devido ao acaso, que deve ser maior ou igual a 80% (vide Quadro 2).

Quadro 2 - Há duas possibilidades para a hipótese nula H_0: ser verdadeira ou falsa. Como para a estatística o resultado de uma pesquisa é sempre devido ao acaso, o pesquisador sempre estará correto em aceitar a H_0. Entretanto, não se faz pesquisa para encontrar resultados por acaso. Por isso, quando a probabilidade de o resultado ter sido por acaso é pequena (≤ 5%), o pesquisador, ao rejeitar a H_0, corre o pequeno risco de estar errado (erro α). Na outra possibilidade, se a H_0 for falsa, o pesquisador estará errado ao aceitar a H_0 (erro β ≤ 20%) e estará correto em rejeitar a H_0 (Poder = 1 – erro β)

Erros intrínsecos da pesquisa científica

Decisão	Hipótese nula	
	Verdadeira	Falsa
Aceitar	Decisão correta	Erro β
Rejeitar	Erro α	Decisão correta (Poder)

Os campos de entrada dos dados para se calcular o poder (1 – erro β) nos nove desenhos de pesquisas apresentados neste livro são os mesmos campos exigidos para calcular o tamanho da amostra. O que poderá mudar serão os dados que alimentam estes campos. Para o cálculo do poder (1 – erro β), o campo ***Power (1 – β)*** é substituído pelo campo ***Total sample size***, ou seja, o verdadeiro tamanho amostral alcançado na pesquisa deve ser informado e não o estimado anteriormente, para verificar se, apesar de uma amostra menor, o tamanho é suficiente para um poder ≥ 0.80. De uma forma geral, enquanto o cálculo amostral se baseou em estimativas do tamanho do efeito, da taxa de alocação dos participantes para os grupos, da correlação entre as medidas repetidas, da odds ratio, da probabilidade de ocorrência da hipótese nula (H_0), da multicolinearidade entre as variáveis preditoras, do aumento relativo da taxa de eventos, da taxa média de eventos assumida em hipótese nula (H_0), da unidade de tempo durante a qual os eventos são contados, a análise do poder (1 – erro β) se baseará nas exatidões destes dados obtidos no final da pesquisa, considerando cada desenho da pesquisa.

Suponhamos que uma pesquisa de dois grupos com três medidas repetidas obteve somente 26 participantes do tamanho amostral estimado de 36 participantes (vide Figura 10). Entretanto, ao final da pesquisa, o tamanho do efeito f foi de 0.30 (e não o de 0.2221308 utilizado no cálculo amostral), o tamanho da amostra foi de 26 participantes (e não o de 36 participantes estimado no cálculo amostral) e a correlação entre as medidas foi de 0.6 (e não a de 0.5 utilizado no cálculo amostral).

Abrindo o aplicativo G*Power, aparecerá o quadro principal com os seguintes campos: ***Test Family; Statistical test; Type of power analysis; Effect size f; α err prob; Total sample size; Number of groups; Number of measruments; Corr among rep measures; Nonsphericity correction Ɛ.*** Preencha estes campos com os dados abaixo e vide a Figura 20:

> ***Test Family*** *= F tests*
>
> A família de testes a ser selecionada é aquela que corresponde ao desenho da pesquisa.
>
> ***Statistical test*** *= ANOVA: Repeated measures, within-between interaction*
>
> O teste estatístico a ser selecionado é aquele correspondente ao desenho da pesquisa e que será utilizado na análise dos dados da pesquisa.
>
> ***Type of power analysis*** *= Post hoc: Compute achieved power – given α, sample size, and effect size*

O poder (1 – erro β) alcançado depende do erro α, do tamanho da amostra ao final da pesquisa e do tamanho do efeito ao final da pesquisa.

Effect size f *= 0.30*

O tamanho do efeito é aquele obtido no final da pesquisa.

α err prob *= 0.05*

Para a amostra obtida no final da pesquisa, 0.05 será o erro máximo aceitável para o pesquisador afirmar que a diferença entre as médias dos grupos não foi devido ao acaso. Erros menores que este vão requerer amostras maiores. Por isto, 0.05 é o padrão a ser utilizado na determinação do poder (1 – erro β).

Total sample size *= 26*

O tamanho da amostra que finalizou a pesquisa.

Number of groups *= 2*

É a quantidade de grupos existente no desenho da pesquisa.

Number of measurements *= 3*

É a quantidade de medidas repetidas do desfecho existente no desenho da pesquisa.

Corr among rep measures *= 0.6*

É a menor correlação da matriz r de Pearson entre as medidas repetidas obtida no final da pesquisa.

Nonsphericity correction Ɛ *= 1*

É a correção da ausência de esfericidade entre as medidas repetidas, podendo variar de 1 / (***Number of measurements*** – 1) ≤ ***Ɛ*** ≤ 1. Quanto menor a correção, maior o tamanho amostral. No cálculo amostral, o valor 1 é o padrão utilizado.

Clique em ***Calculate*** e veja o poder (1 – erro β) no campo ***Power (1 – β err prob)*** = 0.9616630, ou seja, a probabilidade do pesquisador está certo em afirmar que o resultado da pesquisa não foi devido ao acaso é de 96% (vide Figura 20).

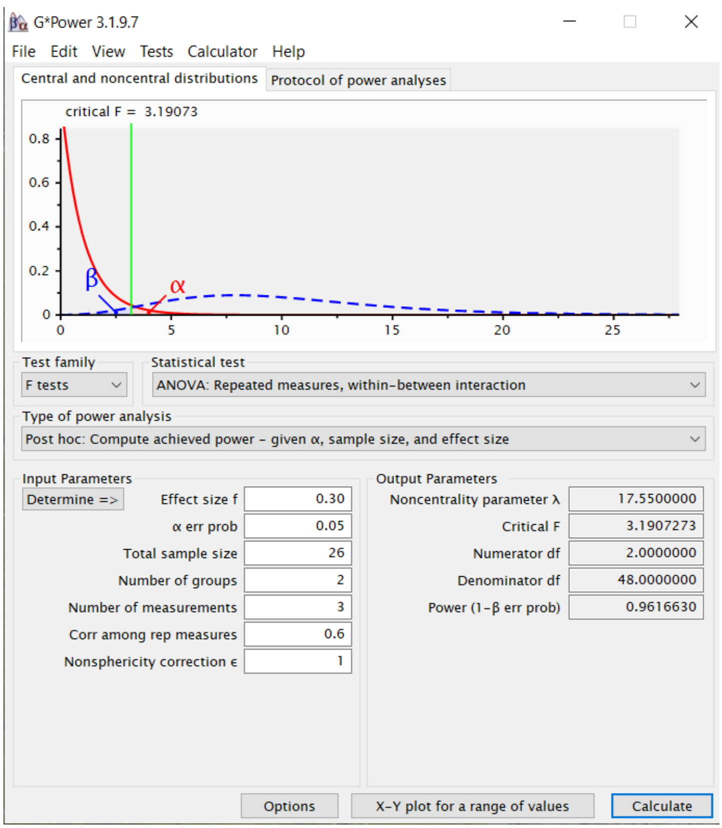


Figura 20 – Determinação do poder (1 – erro β) alcançado de uma pesquisa de dois grupos com três medidas repetidas, utilizando-se o tamanho da amostra, o tamanho do efeito f e a menor correlação r de Pearson entre as medidas repetidas obtidos ao final da pesquisa

Para pesquisa de um grupo e duas medidas repetidas, os dados a serem obtidos no final da pesquisa serão o tamanho do efeito (***Effect size dz***) e o tamanho da amostra (***Total sample size***).

Para pesquisa de dois grupos e uma medida, os dados a serem obtidos no final da pesquisa serão o tamanho do efeito (***Effect size d***), o tamanho da amostra de um grupo (***Sample size group 1***) e o tamanho da amostra do outro grupo (***Sample size group 2***).

Para pesquisa de três ou mais grupos e uma medida, os dados a serem obtidos no final da pesquisa serão o tamanho do efeito (***Effect size f***) e o tamanho da amostra (***Total sample size***).

Para pesquisa de um grupo e três ou mais medidas repetidas, os dados a serem obtidos no final da pesquisa serão o tamanho do efeito (***Effect size f***), o tamanho da amostra (***Total sample size***) e a menor correlação entre a medidas repetidas (***Corr among rep measures***).

Para pesquisa de dois ou mais grupos e duas ou mais medidas repetidas, os dados a serem obtidos no final da pesquisa serão o tamanho do efeito (***Effect size f***), o tamanho da amostra (***Total sample size***) e a menor correlação entre as medidas repetidas (**Corr among rep measures**).

Para pesquisa de correlação entre duas variáveis, os dados a serem obtidos no final da pesquisa serão a correlação r de Pearson (***Correlation*** $\boldsymbol{\rho\ H_1}$) e o tamanho da amostra (***Total sample size***).

Para pesquisa de predição de uma variável contínua intervalar a partir de duas ou mais variáveis, os dados a serem obtidos no final da pesquisa serão o tamanho do efeito (***Effect size*** $\boldsymbol{f^2}$) e o tamanho da amostra (***Total sample size***).

Para pesquisa de predição de uma variável nominal, que expressa uma dicotomia, a partir de duas ou mais variáveis, os dados a serem obtidos no final da pesquisa serão a razão de chances (***Odds ratio***), o tamanho da amostra (***Total sample size***), a multicolinearidade (***R^2 other X***) e o tipo de distribuição das variáveis preditoras (***X distribution***).

Para pesquisa de predição de uma variável nominal, que expressa uma quantidade, a partir de duas ou mais variáveis, os dados a serem obtidos no final da pesquisa serão o aumento relativo da taxa de eventos (***Exp (β_1)***), o tamanho da amostra (***Total sample size***), a taxa média de eventos assumida em hipótese nula (H_0) (***Base rate exp (β_0)***), a unidade de tempo durante a qual os eventos são contados (***Mean exposure***), a multicolinearidade (***R^2 other X***) e o tipo de distribuição das variáveis preditoras (***X distribution***).

Fatores determinantes do cálculo amostral para determinar a média ou proporção de uma variável da população

Fazer adequadamente uma amostragem requer estimar a quantidade necessária de participantes e, esta quantidade estimada chamada de amostra, deve conter as principais características da população da qual está sendo extraída a amostra.

Os fatores determinantes para estimar o tamanho de uma amostra para pesquisa que visa determinar a média ou proporção de uma variável da população são: (1) o nível de confiança; (2) a heterogeneidade da variável na população; (3) e o erro admitido. Se considerados estes fatores determinantes do cálculo amostral, então esta quantidade amostral poderá ser utilizada para determinar a média ou proporção de uma variável da população.

Nível de confiança

O nível de confiança a ser estimado pelo pesquisador pode variar de 0 a 100% e é representado pelo escore z. Escore z é uma estatística que representa o afastamento padronizado de um resultado acima e abaixo do resultado da amostra e de acordo com o tamanho desse afastamento, há um percentual de resultados contidos nos limites desse afastamento (vide Tabela 3). O nível de confiança frequentemente utilizado para o cálculo amostral é de 95%, correspondente ao escore $z = 1{,}96$.

Tabela 3 – Correspondência entre escore z e nível de confiança

Escore z	Nível de confiança
± 3	99%
± 1,96	95%
± 1,65	90%
± 1,28	80%
± 1,04	70%

Heterogeneidade da variável na população

Uma população é considerada homogênea quando a variável a ser estudada está presente ou prevalente na maioria de seus integrantes. O contrário, quando não está presente ou prevalente na maioria de seus integrantes, a população é considerada heterogênea. A heterogeneidade da variável na população é também estimada pelo pesquisador ou pode ser determinada a partir de uma pesquisa semelhante publicada ou de um estudo piloto. O desvio padrão de uma variável contínua intervalar representa sua heterogeneidade, enquanto a prevalência representa a heterogeneidade de uma variável nominal na população.

Erro admitido

É a diferença máxima admitida entre o resultado da amostra e o resultado da população, caso a população pudesse ter sido avaliada.

Geralmente, o tamanho da amostra (n) para determinar a média ou proporção de uma variável da população é dada por

$$n = \frac{\text{Nível de confiança} * \text{Heterogeneidade}}{\text{Erro admitido}}$$

Interpretando a fórmula acima, o tamanho da amostra será proporcional ao nível de confiança estabelecido e à heterogeneidade da variável na população e inversamente proporcional ao erro admitido. Quanto maior o nível de confiança estabelecido, maior será o tamanho da amostra. Quanto maior a heterogeneidade da variável na população, maior será o tamanho da amostra. Quanto menor o erro admitido, maior será o tamanho da amostra.

Há razão para quem observou, à primeira vista, que o cálculo amostral para diferença de médias ou proporções entre grupos ou relação entre duas ou mais variáveis é diferente do cálculo amostral para média ou proporção de uma variável da população. Entretanto, um olhar "falcão peregrino" mostra a mesma racionalidade e correspondência dos fatores determinantes entre as duas situações: Quanto maior o tamanho do efeito ou menor a heterogeneidade da variável na população, menor será o tamanho da amostra; quanto maior o erro α ou o erro admitido, menor será o tamanho da amostra; quanto menor o poder (1 – erro β) ou o nível de confiança, menor será o tamanho da amostra (vide Quadros 3 e 4).

Quadro 3 - Tamanho da amostra estimado em função dos fatores determinantes

Fatores determinantes	Cálculo amostral
Nível de confiança menor	Tamanho da amostra será menor
Heterogeneidade pequena	
Erro admitido maior	

Quadro 4 – Correspondência dos fatores determinantes entre (1) e (2)

Fatores determinantes	
(1) Diferença de médias ou proporções entre grupos ou relação entre duas ou mais variáveis	(2) Média ou proporção de uma variável da população
Tamanho do efeito grande	Heterogeneidade pequena
Erro α maior	Erro admitido maior
Poder (1 – erro β) menor	Nível de confiança menor

Cálculo amostral para determinar a média ou proporção de uma variável da população

Apesar dos fatores determinantes serem comuns para o cálculo amostral para determinar a média ou proporção de uma variável da população, há uma sutil diferença no arranjo destes fatores na constituição da fórmula. Especificamente, o tamanho da amostra (n) para média de uma variável da população pode ser dado por

$$n = \left(\frac{\text{escore } z * \text{desvio padrão}}{\text{erro admitido}}\right)^2$$

e para a proporção de uma variável da população pode ser dado por

$$n = \frac{\text{escore } z^2 * p\,(1 - p)}{\text{erro admitido}^2}$$

Se a população for finita, o tamanho da amostra (n) calculado acima deve ser ajustado à quantidade de indivíduos da população (N):

$$n \text{ ajustado} = \frac{n * N}{(n + N)}$$

Sendo o escore z corresponde ao nível de confiança estabelecido (vide Tabela 3); desvio padrão corresponde a heterogeneidade da variável na população; erro admitido corresponde ao erro entre o resultado da amostra e o resultado da população; e p corresponde a prevalência da variável na população; n o tamanho da amostra estimado e N o tamanho da população.

Tamanho da amostra para determinar a média de uma variável da população

Supondo que o reitor deseje saber a nota média de uma universidade com 20.000 alunos nos três turnos de aula: manhã, tarde e noite. O desvio padrão estimado da nota dos alunos foi de 2 e o erro admitido foi de 4%. Qual seria o tamanho da amostra estimado?

$$n = \left(\frac{\text{escore z} * \text{desvio padrão}}{\text{erro admitido}}\right)^2$$

$$\text{n ajustado} = \frac{n * N}{(n + N)}$$

$$n = \left(\frac{1{,}96 * 2}{0{,}04}\right)^2 = 9.604 \text{ alunos}$$

$$\text{n ajustado} = \frac{2.450 * 20.000}{(2.450 + 20.000)} = 6.488 \text{ alunos}$$

Para que a amostra de 6.488 alunos seja representativa da população de 20.000 alunos daquela universidade, será necessário que a amostra considere as principais características, chamadas também de estratos, dessa população, que possam interferir no resultado da nota média da universidade. Como exemplo, um dos estratos é o turno das aulas: há 10.000 alunos no turno da manhã, 6.000 alunos no turno da tarde e 4.000 alunos no turno da noite, ou seja, dos 20.000 alunos da universidade, 50% são do turno da manhã, 30% são do turno da tarde e 20% são do turno da noite.

Para representar esse estrato de turno das aulas, na amostra de 6.488 alunos, 3.244 (50%) devem ser alunos do turno da manhã, e 1.946 (30%) devem ser alunos do turno da tarde e 1.298 (20%) devem ser alunos do turno da noite. Outros estratos que possam a vir interferir na nota média da universidade devem ser considerados. Por exemplo, a quantidade de alunos em cada graduação. Atender as características da população é tão fundamental quanto a quantidade calculada para uma boa amostragem.

Tamanho da amostra para determinar a proporção de uma variável da população

Supondo que um pesquisador foi contratado por uma instituição pública com 10.000 trabalhadores com diferentes vínculos empregatícios para calcular o tamanho amostral estimado para levantar se esta instituição pública realiza seu trabalho com ética. O presidente da instituição pública pediu um resultado confiável, com pouco erro e o mais rápido possível.

O pesquisador, considerando o pedido do presidente, estabeleceu o nível de confiança de 95%, que corresponde ao escore z = 1,96; como o pesquisador desconhecia a prevalência do trabalho ético naquela instituição e foi imposto rapidez, ele abriu mão do estudo piloto e considerou a menor prevalência de 0,50; e admitiu um erro de 3%.

$$n = \frac{\text{escore } z^2 * p\ (1 - p)}{\text{erro admitido}^2}$$

$$n \text{ ajustado} = \frac{n * N}{(n + N)}$$

$$n = \frac{1{,}96^2 * 0{,}50\ (1 - 0{,}50)}{0{,}03^2} = 1.067 \text{ trabalhadores}$$

$$n \text{ ajustado} = \frac{1.067 * 10.000}{(1.067 + 10.000)} = 964 \text{ trabalhadores}$$

Para que a amostra de 964 participantes seja representativa da população de 10.000 trabalhadores daquela instituição pública, será necessário que a amostra considere as principais características, chamadas também de estratos, dessa população, que possam interferir no resultado da presença de ética no trabalho realizado pela instituição. Como exemplo, um dos estratos é o vínculo empregatício: há na instituição 4.038 trabalhadores efetivos e 5.962 trabalhadores temporários, ou seja, dos 10.000 trabalhadores da instituição pública, 40% são trabalhadores efetivos e 60% são trabalhadores temporários.

Para representar esse estrato de vínculo empregatício, na amostra de 964 participantes, 386 (40%) devem ser trabalhadores efetivos e 578 (60%) devem trabalhadores temporários. Outros estratos que possam a vir interferir no resultado da presença de ética no trabalho realizado pela instituição devem ser considerados. Vou repetir porque é importante: atender as características da população é tão fundamental quanto a quantidade calculada para uma boa amostragem.

Alocação dos participantes para os grupos

Estimado o tamanho da amostra, como fazer a alocação aleatória dos participantes para os grupos? O uso conjunto da função SE com a função ALEATORIO() do aplicativo Excel pode gerar com facilidade eficiência e confiança uma tabela de números aleatórios:

=SE(ALEATÓRIO()<0,500001;1;2) gerará aleatoriamente os números "1" ou "2", para pesquisas com dois grupos;

=SE(ALEATÓRIO()<0,333334;1;SE(ALEATÓRIO()<0,666667;2;SE(ALEATÓRIO()<1,000001;3))) gerará aleatoriamente os números "1", "2" ou "3", para pesquisas com três grupos;

=SE(ALEATÓRIO()<0,250001;1;SE(ALEATÓRIO()<0,500001;2;SE(ALEATÓRIO()<0,750001;3;SE(ALEATÓRIO()<1,000001;4)))) gerará aleatoriamente os números "1", "2", "3" ou "4", para pesquisas com quatro grupos.

Copie os comandos acima e cole na caixa ao lado de "*fx*" no Excel. Arraste o número gerado na casela pela quantidade de caselas correspondente ao tamanho amostral estimado. Os números aleatórios gerados representarão o grupo para o qual será alocado o participante admitido na triagem.

Epílogo

Quem leu este livro até aqui, descobriu que o cálculo amostral para determinar o tamanho de uma amostra de pesquisa, que visa testar a diferença de médias entre grupos ou testar a relação entre duas ou mais variáveis, segue um “algoritmo”:

1) A família de testes e o teste estatístico a serem selecionados são aqueles que correspondem ao desenho da pesquisa;
2) O erro α = 0.05, porque é o erro máximo aceitável;
3) O poder (1 – erro β) = 0.80, porque é o acerto mínimo aceitável;
4) O tamanho do efeito é o pequeno, correspondente ao teste estatístico que será utilizado na pesquisa;
5) Se o tamanho da amostra for muito grande, difícil de ser obtido, outro tamanho do efeito, obtido a partir dos resultados de um estudo piloto ou pesquisa semelhante publicada, deve ser determinado e utilizado para o cálculo amostral.

Da mesma forma, a determinação do poder (1 – erro β) de uma pesquisa que não obteve o tamanho da amostra estimado, que visou testar a diferença de médias entre grupos, segue também um “algoritmo”:

1) A família de testes e o teste estatístico a serem selecionados são aqueles que correspondem ao desenho da pesquisa;
2) O erro α = 0.05, porque é o erro máximo aceitável;
3) O campo ***Power (1 – β)*** é substituído pelo campo ***Total sample size***, ou seja, o verdadeiro tamanho amostral alcançado na pesquisa deve ser informado e não o estimado anteriormente, para ver se, apesar de uma amostra menor, o tamanho é suficiente para um poder ≥ 0.80;

4) Os dados do tamanho do efeito, da taxa de alocação dos participantes para os grupos e da menor correlação entre as medidas repetidas são aqueles obtidos no final da pesquisa.

Da mesma maneira, a determinação do poder (1 – erro β) de uma pesquisa que não obteve o tamanho da amostra estimado, que visou testar a relação entre duas ou mais variáveis, segue também um “algoritmo”:

1) A família de testes e o teste estatístico a serem selecionados são aqueles que correspondem ao desenho da pesquisa;
2) O erro $\alpha = 0.05$, porque é o erro máximo aceitável;
3) O campo ***Power (1 – β)*** é substituído pelo campo ***Total sample size***, ou seja, o verdadeiro tamanho amostral alcançado na pesquisa deve ser informado e não o estimado anteriormente, para ver se, apesar de uma amostra menor, o tamanho é suficiente para um poder ≥ 0.80;
4) Os dados da *odds ratio*, da probabilidade de ocorrência da hipótese nula (H_0), da multicolinearidade entre as variáveis preditoras, do aumento relativo da taxa de eventos, da taxa média de eventos assumida em hipótese nula (H_0), da unidade de tempo durante a qual os eventos são contados são aqueles obtidos no final da pesquisa, considerando cada desenho da pesquisa.

Assim, o cálculo amostral para determinar o tamanho de uma amostra para pesquisas que visem determinar a média ou proporção de uma variável da população segue também um algoritmo:

1) Informe o nível de confiança pretendido (escore Z);
2) Informe a heterogeneidade da variável na população (desvio padrão para variável contínua e prevalência para variável proporção);

3) Informe o erro admitido.

Em suma, é importante destacar que o cálculo amostral é uma estimativa da quantidade de participantes necessária para uma pesquisa. Sendo estimativa, contém erro. Entretanto, este erro é muito menor do que aquele advindo de uma estimativa arbitrária. O erro que se pode cometer no cálculo amostral é inversamente proporcional ao conhecimento do pesquisador sobre estatística, metodologia da pesquisa e o assunto a ser pesquisado.

Bibliografia

Altman DG, et al. Concealing treatment allocation in randomized trials. BMJ. 2001;323(7310):446-7. DOI: 10.1136/bmj.323.7310.446.

Hopkins WG. A new view of statistics [Internet]. 2016 [Acesso em: 21 set 2024]. Disponível em: https://www.sportsci.org/resource/stats.

Bruni AL. Estatística aplicada à gestão empresarial. 4 ed. São Paulo: Atlas; 2013. 416 p.

Buchner A, et al. G* Power 3.1 manual [Internet]. 2023. [Acesso em: 30 ago 2024]. Disponível em: https://www.kent.ac.uk/software/gpower.

Carneiro AV. Cálculo da dimensão da amostra em estudos clínicos: princípios metodológicos básicos. Rev Port Cardiol. 2003;22(12):1513-21. Disponível em: https://repositorio.ulisboa.pt/bitstream/10451/33057/1/Calculo_dimensao_amostra.pdf.

Cohen J. A power primer. Psychol Bull. 1992;112(1):155-9. DOI: 10.1037//0033-2909.112.1.155.

Cohen J. Statistical power analysis. Curr Dir Psychol Sci. 1992;1(3):98-101. DOI: 10.1111/1467-8721.ep10768783.

Ellis PD. The essential guide to effect sizes. 1 ed. Cambridge: Cambridge University Press; 2010. 188 p.

Field A. Descobrindo a estatística usando o SPSS. 2 ed. Penso; 2009. 269 p.

Godwin M. Hypothesis: the research page. Part 3: Power, sample size, and clinical significance. Can Fam Physician. 2001;47:1441-3, 1450-3. PMID: 11494932, PMCID: PMC2018536.

Gordon G, et al. Grading of Recommendations Assessment, Development and Evaluation (GRADE) [Internet]. 2000. [Acesso em: 10 ago 2024].Disponível em: https://www.gradeworkinggroup.org/.

Hewitt CE, et al. Is restricted randomization necessary? BMJ. 2006;332(7556):1506-8. DOI: 10.1136/bmj.332.7556.1506.

Paul G, Iain C. Is 85% of health research really "wasted"? [Internet]. 2016 [Acesso em: 24 out 2024]. Disponível em: https://blogs.bmj.com/bmj/2016/01/14/paul-glasziou-and-iain-chalmers-is-85-of-health-research-really-wasted/.

Jadad AR, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996;17(1):1-12. DOI: 10.1016/0197-2456(95)00134-4.

Jones SR, et al. An introduction to power and sample size estimation. Emerg Med J. 2003;20(5):453-8. DOI: 10.1136/emj.20.5.453.

Lachin JM, et al. Randomization in clinical trials: conclusion and recommendations. Control Clin Trials. 1988;9(4):365-74. DOI: 10.1016/0197-2456(88)90049-9.

Marotti J, et al. Amostragem em pesquisa clínica: tamanho da amostra. Rev Odontol Univ Cidade S Paulo. 2008;20(2):186-94. Disponível em: https://www.researchgate.net/publication/285800533_Amostragem_em_pesquisa_clinica_Tamanho_da_amostra.

Mendes M, et al. Type I Error Rate and Power of Three Normality Tests. Pak J Inf Technol. 2003;2(2):135-9. Disponível em: https://www.researchgate.net/publication/26556526_Type_I_Error_Rate_and_Power_of_Three_Normality_Tests.

Rubini EC, et al. Critical analysis of the drafting of physical therapy randomized controlled trials published in Portuguese. Fisioter Mov. 2016;29(2):421-7. DOI: 10.1590/0103-5150.029.002.AO21.

Schulz KF. Assessing allocation concealment and blinding in randomised controlled trials: why bother? Evid Based Nurs. 2001;4(1):4-6. DOI: 10.1136/ebn.4.1.4.

Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet. 2002;359(9305):515-9. DOI: 10.1016/S0140-6736(02)07683-3.

Thomas JR. Métodos de pesquisa em atividade física. 6 ed. Porto Alegre: Artmed; 2012. 478 p.

Torgerson DJ, Roberts C. Understanding controlled trials. Randomisation methods: concealment. BMJ. 1999;319(7206):375-6. DOI: 10.1136/bmj.319.7206.375.

Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010; 36(3): 1-48. DOI: 10.18637/jss.v036.i03

Vincent WJ. Statistics in kinesiology. 2 ed. Champaign: Human Kinetics; 2005. 293 p.

Azar daquele que não
sabe a sorte que tem.

Autor desconhecido

Pedro & João Editores
20 Anos
pedroejoaoeditores.com.br

ISBN 978-65-265-1880-9