**Proposta de orientação sobre cálculo de tamanho amostral**

Cálculos de tamanho amostral e poder estatístico

Seguindo a tendência de diversos órgãos regulamentadores ao redor do mundo, o CEUA-CCS exige a realização de cálculos de poder estatístico e tamanho amostral para a autorização de experimentos em animais. A razão é simples: experimentos com baixo poder estatístico tem uma grande chance tanto de apresentarem resultados falsamente negativos (pela própria definição do conceito de poder estatístico, que é a capacidade de um teste estatístico detectar um determinado tamanho de efeito se ele existir) como de resultados positivos (já que, em se utilizando um limiar de significância fixo, um resultado significativo terá menos chance de ser verdadeiro se o poder estatístico for baixo). Desta forma, a ausência de um cálculo de tamanho amostral e poder estatístico representam a exposição de animais de experimentação a experimentos inapropriados para responder uma pergunta experimental.

A realização de cálculos de poder estatístico para a maior parte dos experimentos a serem analisados por testes de estatística paramétrica, é simples e pode ser realizada em inúmeras calculadoras online (como em <http://powerandsamplesize.com> e <http://www.stat.ubc.ca/~rollin/stats/ssize/>) ou softwares disponíveis gratuitamente (como o GPower, disponível em <http://www.gpower.hhu.de/>).

Em resumo, 4 parâmetros básicos são necessários para calcular o tamanho de amostra recomendado para um experimento:

1. O tamanho da diferença entre grupos que se planeja detectar com o estudo (o qual deve ser estabelecido pelo pesquisador).

2. O poder estatístico desejado (recomendamos que o mesmo seja de pelo menos 80%).

3. A variância ou desvio padrão esperado para os grupos experimentais.

4. O limiar de significância a ser utilizado no teste estatístico.

Diferentes testes paramétricos ou calculadoras podem requerer que as informações (e.g. a definição de tamanho de diferença) sejam descritos de maneira discretamente diferente. Para ilustrar o processo com um exemplo simples, vamos mostrar como os parâmetros devem ser utilizados para um cálculo de tamanho amostral de um teste t bicaudal para amostras independentes no site <http://powerandsamplesize.com>:

1. Entre no site

2. Clique em “go straight to the calculators” e em “Compare 2 means – 2 samples – 2 sided equality” (isso levará você para este link: <http://powerandsamplesize.com/Calculators/Compare-2-Means/2-Sample-Equality>).

3. No primeiro campo de cima para baixo, você pode optar em calcular (a) um tamanho amostral para obter um determinado poder estatístico ou (b) o poder estatístico de um determinado tamanho amostral. Para fins de planejamento experimental, o caso (a) será mais comum, então vamos analisá-lo a seguir.

4. Na linha abaixo, você deve inserir (a) o poder estatístico desejado (em forma de decimal – e.g. “0.8” significa 80%) e (b) a chance do erro tipo 1, que equivale ao limiar de significância estatística desejado (e.g. 5% para um valor p de 0.05). O campo “sample size” não deve ser preenchido, pois é exatamente este número que vamos calcular em nosso exemplo.

5. Nas próximas linhas (“Group ‘A’ mean” e “Group ‘B’ mean) você deve especificar o tamanho de diferença que pretende detectar, ao inserir as médias esperadas para os dois grupos no caso desta diferença existir. Por exemplo, se você está desenhando um experimento capaz de detectar uma redução de 50% em um determinado parâmetro em relação ao grupo controle (cuja média esperada é de 40 mg/dl, por exemplo), pode inserir 40 no grupo A (controle) e 20 no grupo B (tratado). Note que as unidades nesse caso não importam (e.g. você poderia inserir “100” e “50” se quisesse considerar o valor do controle como 100%), mas lembre que *o desvio padrão deverá ser inserido na mesma unidade que as médias*.

7. Na linha seguinte, você deverá inserir o desvio padrão. No caso do teste t, o mesmo usa um desvio padrão calculado a partir dos desvios dos dois grupos (calculado por  s_{X_1X_2} = \sqrt{(s_{X_1}^2+s_{X_2}^2)}). Se você não dispuser de um desvio padrão esperado para o grupo tratado (uma situação comum), utilizar a estimativa de desvio padrão dos controles é uma alternativa razoável, a não ser que seu tratamento aumente de forma significativa a variância da variável a ser medida entre os indivíduos. Novamente, lembre que o desvio padrão deve estar na mesma unidade que as médias – i.e. se você espera uma média de 40 mg/dl, um desvio padrão de 10 mg/dl e quer detectar uma redução de 50%, você pode inserir tudo em unidades brutas (médias de 40 e 20, desvio padrão de 10) ou normalizadas (médias de 100 e 50 , desvio padrão de 25) sem que isso mude o resultado do teste.

8. A última linha (sampling ratio) contempla a possibilidade de experimentos com grupos de tamanho diferente. Idealmente, o melhor poder estatístico de um estudo sempre será obtido com grupos experimentais do mesmo tamanho (nesse caso, use 1 neste campo). Entretanto, se por alguma razão isso não for factível, você pode alterar o parâmetro inserindo o valor de kappa (definido como a razão entre o tamanho amostral dos dois grupos – i.e. n do grupo A/n do grupo B). Neste caso, o resultado dado pelo teste no campo sample size representará o tamanho amostral necessário para o grupo B, e você deverá calcular o tamanho amostral do grupo A a partir desta razão.

9. Depois de preencher todos os itens, clique em “Calculate” e veja o número que aparece no campo “Sample Size”.

Tomando um exemplo simples, se você tem um experimento planejado para (a) detectar uma redução de 50% em um parâmetro cuja média nos controles é 40 mg/dl, (b) com um desvio padrão de 10 mg/dl, (c) com um poder estatístico de 80%, (d) um limiar de significância de p=0.05, (e) com grupos de tamanho igual, seu poder estatístico será de 4 animais por grupo.

Note que para detectar uma diferença menor (e.g. uma redução de 25%) com o mesmo poder estatístico, o tamanho amostral será maior (nesse caso, 16). Caso o poder estatístico desejado seja maior (e.g. 90%), o tamanho amostral também aumenta (no caso anterior, de 4 para 6).

Cálculos de outros testes paramétricos, como testes t de 1 amostra, ANOVA de uma via, e comparações entre proporções (e.g. qui-quadrado) também são relativamente simples, e podem ser realizados no próprio site. Cálculos de tamanho amostral para métodos não-paramétricos (como os testes de Mann-Whitney e Kruskal-Wallis) são mais complexos e geralmente dependem de uma estimativa prévia de distribuição, o que pode ser mais difícil de realizar. Nestes casos, uma regra simples é calcular o tamanho amostral para uma comparação paramétrica e adicionar 15%, o que representa uma estimativa conservadora do tamanho amostral necessário.

Para mais informações, qualquer livro de bioestatística possui uma seção de cálculo de tamanho amostral. Recomendamos ainda as seguintes referências:

**XXXXXXXXX**

**XXXXXXXXX**

Atenciosamente,

CEUA- CCS