

Q1. 疾病コーディング分析や、がん登録分析を行う際に注意すべき点や、院内で活用する際に注意すべきことをご教授願います。また、分析する際に使うと便利なツールなどがあればご教授願います。

＜回答＞

（1） 疾病コーディング分析について

疾病コーディング分析とは、院内での疾病統計、たとえば、どの疾患がどれくらい受診したか、または、入院したか？また、その経時的な変化に特徴はあるか？などの分析を指すでしょう。このような分析を行う際には、ハンズオンセミナーでも示しましたが、「分布の特徴を適切な方法でグラフや表にまとめ、その特徴を正しい統計量で表現する」ことだと思います。ヒストグラム、棒グラフ、円グラフ、百分率帯グラフ、折れ線グラフ、箱ひげ図など、データの種類（量的か質的か）によって使い分けが必要です。また、量的データの場合、正規分布か否かによって特徴づけの統計量が異なります。

（2） がん登録分析について

特別セッションで説明いたしましたが、がん登録で集積される生存時間データは特殊なデータです。死亡症例と打ち切り症例（最終生存確認日まで生存していたことはわかっているが、その後いつ死亡したかわからない症例）が混在しますので、特別な統計処理が必要です。がん登録データの分析では、生存率が算出されます。これは Kaplan-Meier 法で算出されますが、Excel では処理できません。以下に述べる統計解析ソフトウェアが有用かと思います。

（3） 院内で活用する際の注意点

疾病統計の報告書や病院管理・病院経営の会議資料の作成などで、統計解析した結果は様々な使われ方をします。注意点を以下にまとめました。

- ・データを集約するグラフや表は見やすいものに仕上げる。
- ・作成されたグラフ等には、その分布の特徴を示すための統計量（平均士標準偏差、中央値（25%点、75%点）、割合、生存率など）を正しく表示する。
- ・得られた図表は統計量の解釈を1～2行で記述する。

（4） 分析に便利なツールについて

Excel は図表の作成をきれいに仕上げるのに適していますが、処理に時間がかかること、少し高度な統計処理ができないことが問題となります。そこで、無料で入手できる対話型統計解析ソフトウェアを利用することをおすすめします。具体的には、EzR（自治医科大学付属さいたま医療センター血液科の HP からダウンロードできます）が図表の作成や医療で使う統計解析処理を行う際に便利かと思います。EzR ダウンロードで検索すれば、ダウンロー

ドの方法が提示されます。

Q2. 統計解析を今までやったことがないのですが、初めて解析する場合のおすすめのデータや解析方法等はありますか？

＜回答＞

初めてですと何から手を付けてよいのかわかりませんよね。今回、ハンズオンセミナーで用意しました二つのデータセットと演習課題を 1 つずつこなしてみてはいかがでしょうか？診療情報管理士にとって実践的な統計処理の方法が網羅されていると思います。解析方法については特別セッションでお話ししましたが、解析目的と処理するデータの種類によって違ってきます。その使い分けが難しいかもしれません、覚えるより慣れていくとよいと思います。そしてわからないことがあれば、ひとりで悩まずに人に尋ねることです。

Q3. パワーピボットが今、一つ使いこなせないため、講義の中に組み込んでほしい。

＜回答＞

パワーピボットについては中級者から上級者向けと考え、セミナーでは取り上げませんでした。パワーピボットの使用例 (PowerQuery でデータを結合する方法) を別添のスライドに用意しました。今回のデモデータ作成のときにも使っております。ただ、使用目的によって使用方法は様々でしょうから、個別のご相談にて対応させていただきたいと思います。

Q4. ピポットテーブルの使い方

＜回答＞

ハンズオンセミナーで典型的なピボットテーブルの使用方法を説明し実習いただきました。もし、他の使用目的でご不明な点があればお尋ねください。

Q5. 例えば「診療単価が高い」というときの定義の仕方について、統計上に考え方があるのであればそれをご教示いただきたいです。

＜回答＞

「診療単価が高い」の定義は、おそらくですが、医療経済・病院経営的にも統計学的にも公式なものはないと考えます。診療単価が高いか否かは、他の医療施設と比較して「相対的に高い」を意味するものと考えます。すなわち、ベンチマーク分析です。もし、疾患別や病院別の診療単価の生データかもしれませんのはその分布を特徴づける平均と標準偏差などが公表されていれば、自院が

全国の病院の中でどれくらいの位置を占めるか（診療単価が平均より高いか低いか、標準偏差で見てどれくらい高いかなど）を評価するものと考えます。実際に、国立大学法人の病院では、病院経営のデータを提出しあってそれぞれの病院の相対的評価を行っています。自院だけのデータで診療単価の高い・低いを議論するのであれば、前年度に比べてみるという方法もあるかと思います。

Q6. 院内がん登録データを用いて、がん患者居住地分布状況を可視化できたらと考えております。Excelで作成することはできますでしょうか。

＜回答＞

Excelのマップ機能を使用した例を別添のスライドにまとめてみました。「がん登録の居住地分布の可視化」の例です。使用目的によってさまざまな使用方法が出てくると思いますので、個別の相談にも応じたいと思います。

Q7. 統計解析で得られた結果を、院内での業務改善や意思決定にどう活用すればよいでしょうか？例えば、どのような部署や会議体に共有し、どういったアクションにつなげるのが効果的か、具体例があれば教えていただけますと幸いです。よろしくお願ひいたします。

Q8. 國際疾病統計の基本的な考え方とその手法について

＜回答＞

国際疾病統計が何を指すのか、すみません、私の不勉強で理解できませんでした。ICDを使った様々な統計処理の結果だと思うのですが、使用目的によって考え方や解析手法が変わってきます。もう少し、具体的な質問にしただければお答えできたかもしれません。

Q9. 結果、過剰査定されてしまった検査や注射の件数や金額

＜回答＞

過剰査定された検査や注射はその種別ごとに診療科名や査定額などを付記した形で、毎月病院の運営会議などに出てきます。種類がそれほど多くなく件数もある程度多いのであれば、グラフ化して資料に添えるとよりわかりやすいのではないでしょうか？それと、経時的に折れ線グラフや表で表すと対策の効果が把握しやすいと考えます。

Q10. ①研究目的や状況に応じた適切な統計手法の選択基準

②学会発表に値する研究事例を見出すための視点や着眼点や、分析の計画か

ら実施、結果の解釈に至るまでの一連の手順について実践的なワークフロー やチェックポイントなど。

＜回答＞

①について

これは多種多様で簡単にはお答えしかねます。研究では特別セッションでお話しした推測統計を使うことになりますが、データの分布の特徴づけを数値で表す推定については特別セッションでお話しした記述統計の手法と同じ選択基準となります。問題は検定ですが、こちらの方は解析目的、データの種類、比較する群の数、症例数、正規分布の仮定などにより数十種類の検定手法があります。個別の相談ということにさせてください。

②について

これも多種多様で一概にマニュアル化、標準化できるものではありません。統計解析を行い論文化するための手順を記述したテキストもありますので、それを購読されるのも良いかと思います。この質問に対する回答は、ケースバイケースで異なりますので、個別の相談とさせてください。

Q11. 時系列グラフで年数ごとの比較に中央値を用いてもよいでしょうか。中央値が外れ値を反映しないのは承知していますが、中央値や平均値を用いて良いか疑問になりました。ご教示いただけすると幸いです。

＜回答＞

時系列グラフを折れ線グラフではなく、年次ごとの箱ひげ図で示すこともあります。つまり、中央値の経時的な変化とデータのばらつきの挙動を把握することができます。各年次の量的データが正規分布である場合には平均と標準偏差、正規分布から外れていたり、外れ値が複数存在する場合には、中央値と四分位数を使うことが原則となります。

Q12. 分析ツールでどんなことができるか知りたい。特に、基本統計量以外に、何ができるのは知りたい。

＜回答＞

分析ツールは様々な統計解析を対話型で処理できる優れものです。ハンズオンセミナーでは基本統計量の算出と分散が等しくない場合の t 検定を行いました。これ以外にも、2 群の母分散が等しいかを検定する F 検定、3 群の母平均の差の検定を行う分散分析、2 つの量的データの関連の強さを数値化する相関分析、説明変数から目的変数の値を予測する回帰分析などが可能です。

Q13. どうしたら統計を好きになれますか？

＜回答＞

統計学のテキストを読んでいても数式が多く出てきたり、難解な用語が出てきたりでわかりにくいですよね。まずは、データを分析してみることが初めてはいかがでしょうか？ハンズオンセミナーで実習したヒストグラムの作成や箱ひげ図の作成、クロス集計表の作成など、自分がやってみたいところから出力してみることです。できれば、デモデータを見ながら自分で解析目的を考えてそれをグラフ化してみて結果を解釈する、という実践を積むことが好きになるコツかもしれません。それと、わからない点は遠慮せずに人に聞くことです。

Q14. 統計学というと漠然としていて、どうしたらいいのか分かりません。基本的なことから学びたいです。

＜回答＞

そうですね、この質問の回答もQ13の回答と同じかもしれません。まずは実践的にいろいろ手を動かして出力してみる、でいかがでしょうか？いろいろ出力しているうちに「あれ？」と思うことがあれば、それをまた再解析してみるのです。デモデータA、デモデータBは全国の病院のデータですが、病院を患者や診療科、疾患に置き換えれば院内の疾病統計、病院経営、がん登録、DPC分析などに利用できると思います。わからないことは遠慮なくご相談ください。