

1. 平方根のリスト、あると助かります。

Moodle の今週の回のところにアップロードしてあります。ご参考まで。

2. $\frac{\bar{x}-x}{u/\sqrt{n-1}}$ とか $\frac{1}{n-1}(\dots)$ とかの $n-1$ で、どうして 1 引くのがわかりません!

母分散 σ^2 が未知のときの母平均の推定や検定の際には、不偏分散 U^2 を使います。この不偏分散 U を求めるのに $n-1$ を用いるためです。ちなみに、質問文にある前半の式は $\frac{\bar{x}-x}{u/\sqrt{n}}$ または $\frac{\bar{x}-x}{s/\sqrt{n-1}}$ のはず(教科書 p.147 参照)。

なお、 $t = \frac{\bar{x}-x}{u/\sqrt{n}}$ を標本分散 s^2 を使った式に変形すると $t = \frac{\bar{x}-x}{s/\sqrt{n-1}}$ とかける。これは母分散が不明な場合には、標本分散からそのまま統計検定量を計算するのに便利な式となる。

いずれにしろ、統計の課題を解くにあたっては、母分散が分かっているか、それともそれが不明なので標本分散(または不偏分散)を用いないといけないのかを意識しましょう。

3. p.168 「これらの母比率を p_1, p_2, \dots, p_k とするとき, 「これらの母比率とは $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_k$ である」の箇所が激しく理解不能。説明が欲しい。

たしかに、母集団と標本の関係、そして、その母比率、標本比率の関係を意識しないと分からないために分かりづらくなっていると思うので、もう一度解説します。

4. 課題の問題が理解できない。

課題の問題文は分かりづらいのですが、基本的にはその回の講義で扱った内容なのは間違いないので、少なくとも該当する検定法の仮説にどのようにあてはめればよいか考えて見ましょう。標本はどれでどの平均を比較すればよいのか等、数式ではなく日本語で表現されてしまっていて分かりづらく

なっている箇所もある点をご容赦ください。

5. H_0 と H_1 の設定がよくわかりません...

標本はどれで何の統計値を比較すればよいのか、から考えてみましょう。そして、前回にコメントしたとおり、 H_0 に関してはかならず $a =$ という形になる点も気をつけましょう。

さらに、検定する対象となる統計値が分かれば、あとは検定の種類ごとにある程度固定されます。例えば、母平均の検定の場合、帰無仮説 $H_0: \mu = m$ (母平均は特定の値 m である) で決まりますし、それに対応する対立仮説は $H_1: \mu \neq m$ または $\mu < m, \mu > m$ のいずれかから選ぶことになります。問題文の設定がどれを求めているか、よく読んで考えるようにしてください。

6. 試験の日にインフルエンザなどにかかって、やむを得ず休んだ場合は、単位が出なくなりますか? 救済措置はありますか?

急病や交通機関の乱れ等、本人に責の無い理由による休みや遅刻に関しては、一定の配慮をします。理由が明らかになった場合には、迅速に連絡をお願いします。

7. p.161 のやり方で第 8 回課題 B) を解いてしまいました...

問題ありません。近似手順が違うだけの問題ですので、いずれも正解です。基本的には、標本サイズが大きいので正規分布とみなして母分散既知として解いただけですので、 t 分布として解いても答えはよほどのことが無い限り変わらないはずです。

8. 理論値の求め方がよく分かりません。

おそらく独立性の検定と思いますが、全体の度数の比になっているという仮定のもとに、周辺の比率にあわせて、その対象標本群の数にあわせた比をかけているだけです。念のため、もう一度解説を加えたいと思い

ます。

9. χ^2 検定において、理論度数が 5 未満であるクラスは隣接するクラスと合併することですが、血液型のようなデータを扱う際の合併の基準がよく分かりません。A : B : O : AB = 4 : 3 : 2 : 1 というように、比率がわかっているならば、比率の値が近いクラスに合併すればよいのでしょうか。それでも他の血液型と合併することに抵抗はありますが。

はい、抵抗はありますが、 χ^2 分布の性質上的ことですので、しょうがありません。自分で標本数が決められるときは十分な標本数を確保することにより、こうなることを避けるのが通例です。

10. 教科書 p.178 の

$$\text{自由度} = 6, \chi_6^2(0.05) = 12.5$$

したがって、きわどくも

$$t = 13.7 > 12.6 = \chi_6^2(0.05)$$

となっている箇所。12.5 と 12.6 と、値が違っているのはなぜでしょうか？ よく分からなかったです。

教科書の誤植と思われる。元の χ^2 分布表を見ると 12.59 です。

$$\times \text{自由度} = 6, \chi_6^2(0.05) = 12.5$$

$$\text{自由度} = 6, \chi_6^2(0.05) = 12.6$$

11. 実社会において、検定が役立っている具体例があれば教えて欲しいです。論文等で使われるイメージですが、社会に還元できているのでしょうか。

もちろん、科学論文で使われることが多いですが、科学技術がそのまま適用される領域では非常に重要です。医学などは“Evidence-based medicine” という用語により、臨床研究の結果が統計的に有意に優れていると示されない薬や治療法は却下されるように

なっています。また、初回の講義で少し解説したように「ビッグデータ」といったキーワードのもと、統計的なデータ解析をそのままビジネス応用に用いる動きもあるようです。また、世論調査や選挙速報等、リアルタイムで解析したいという要望も多くなってきており、一昔前よりも直接的に統計解析を行うシーンは多くなりつつあります。

感想その他

12. 先日、医学系の TV 番組を見ていて「統計からタミフルが児童の異常行動に作用しているとはいえない」というものをみたのですが、これも統計的仮説検定の一つですよ？「棄却できない」という概念がこれでピンと来た気がしました。

13. 試験も近いので勉強をがんばろうと思った。

14. 記号がいっぱい出てきて、どれが何かが分からなくなりそう…。 μ とか σ とか…。

記号の使い方はある程度固定されています。たとえば、 μ (ミュー) は母平均、 σ (シグマ) は標準偏差など。ただし、文脈によって呼び方が変わっているので注意 (母平均とただの平均が同じものを指している場合など)。教科書をみながら確認しましょう。