

尿路結石発症率の調査、及びベイズ推定

工学部 情報システム工学科 3年 007

上田 拓也

(指導教員：藤崎礼志)

1. 背景・概要

「尿路結石」という疾患の発症率について、主にベイズ統計学を用いて、考察等を行った。

尿路結石とは、尿路系に石が析出し、またそれが詰まってしまう疾患をいう。

2. 研究

2.1.原因の調査

主にインターネットを用いて、尿路結石とはどのような疾患であるか、また、これの発症に影響を与える因子を調査した。

2.2.データ収集

2.1の調査を元に、質問紙を作成し、Web上でアンケート調査を行った。

2.3.データ分析

2.2で収集したデータを用いて、結石発症率のベイズ推定を行った。

○ベイズ推定とは

【定義】

$P(A)$: ある事象 A の発生確率

$P(X)$: ある事象 X の発生確率

$P(A|X)$: 事象 X が発生した条件の下での事象 A の発生確率

としたとき、ベイズの定理により、

$$P(A|X) = \frac{P(X|A)P(A)}{P(X)}$$

と表される。

また、複数条件におけるベイズ推定は、

$P(A')$: ある事象 A' の発生確率

$P(X'), P(Y'), P(Z'), \dots$: 事象 $X', Y', Z', \dots$ の発生確率

$P(A'|X'), P(A'|Y'), P(A'|Z'), \dots$: 事象 $X', Y', Z', \dots$ が発生した条件下での事象 A' の発生確率

としたとき、

$$P(A|X', Y', Z', \dots) = \frac{P(X|A)P(Y|A)P(Z|A) \dots P(A)}{P(X')P(Y')P(Z')}$$

と表される。

上式の右辺は、一般的に統計的に既知の数値であり、本研究では、2.2で得られたデータより、それぞれの発生確率を求めている。

2.4.データ応用

2.3のベイズ推定を用いて、C言語環境で「腹痛発生率テスト」なるものを作成した。（「尿路結石発症率テスト」を作成する予定であったが、データ量の問題、また、尿路結石発症時は大抵の場合、腹痛を伴うことから、「腹痛発生率」で代用した）
こちらにも同様に「はい」、「いいえ」で回答し、これらの回答を条件としたベイズ推定を行うプログラムとなっている。

3. まとめ

本研究では尿路結石の発症率の推定を行ったが、元となるデータ量が少なく、また、アンケートにおける質問の設定が若干曖昧であった事もあり、余り信憑性の高くない結果となってしまった。しかしながら、本テーマは我々の日常生活と密接に関わる問題であり、今後の生活習慣の改善、維持等に、少しでも貢献できればと思う。