

※第 8 回, 第 10 回の演習問題について説明する他, 下記の問題からいくつか選んで説明します。

1. 次の問 1,2 に答えよ。(昨年試験問題)

1. 「航空機事故は続けて起こりやすい」と, 一般によく言われる.
 - (a) ある瞬間に地球上のどこかで航空機事故が起きる危険の度合いは, 時刻によらず常に一定であると仮定する. このとき, ある時刻から次に航空機事故が起きるまでの経過時間は, どのような確率分布にしたがうと考えられるか. この講義で説明した知識にもとづいて論じよ.
 - (b) 上の仮定が正しいとするとき, 「航空機事故は続けて起こりやすい」という言葉が妥当かどうかについて論じよ.
2. 第 2 次世界大戦中, ドイツ軍では, 暗号をイギリス軍に解読されるのを防ぐため, 通信を暗号化する鍵となるコード (いくつかの数字の列) を毎日変更していた. その際, 可能な数字の組み合わせの中からランダムに数字を選び, さらにその月に一度使ったコードは月内には 2 度と使わないことにしていた. この方法は, コードをランダムにして暗号を見破られにくくする方法として正しいか.

2. 次の各項は, 統計学の観点からみて正しい (あるいは適切である) かどうかを答えよ. 正しくない (適切でない) ときは, どのような点がどのように正しくない (適切でない) かを説明せよ.

1. 高校のクラスの卒業生の, 現在の平均年収を調査するため, クラス会幹事が, クラス会のパーティの案内の返信はがきに年収を書く欄をつくり, 回答を求めた.
2. ある年の総選挙で選ばれた衆議院議員の, 血液型の分布を調べたところ, その分布は, 日本人全体の分布からの有意な偏りがあった. したがって, 「血液型と『国会議員になる資質』の間には関連がある」といえる.
3. A 社の電球の平均寿命は, B 社の電球の平均寿命よりも長い. したがって, 価格, 消費電力, 光量など, 寿命以外の条件が全く同じならば, A 社の電球を使うほうがよい.
4. 第 2 次大戦時, イギリス軍の爆撃機は, ドイツ軍の迎撃戦闘機の攻撃により大きな損害を受けた. そこで, イギリス軍では, 帰還した爆撃機乗務員に「機体のどちらの方向から攻撃を受けたか」を聴取し, その方向の防備を厚くするなどの対策を練った.
5. 母集団の平均の区間推定を行なう問題で, 取り出した標本を使って計算した結果, 「母平均の 95 パーセント信頼区間は 20 から 30 である」という結論を得た. このとき, 母平均が 20 以上 30 以下である確率は 95 パーセントである.
6. 全国の食料品店から, いくつかの店を無作為抽出して調査したところ, 全品目のうち 3% について, 内容量が表示されている量に足りない, という不正が発見された. したがって, 消費者は平均して 3% 損をしていることになる.

解答例

1.

1. (a) 「ある瞬間に事故が起きる危険の度合い」は、講義第7回のハザード関数に相当する。問題の条件では、ハザード関数は時刻によらず一定という条件なので、ある時刻から次に事故が起きるまでの待ち時間は、指数分布にしたがう。
(b) 生存関数のある時刻での値は、「その時刻までに事故が起きていない確率」と考えることができる。指数分布では、その関数は時刻を t とするとき $e^{-\lambda t}$ ($\lambda > 0$) であるから、時間が経つにつれ急激に減少する。すなわち、次の事故が起きるまでの時間は、短いほどその可能性は高く、「事故は続けておきやすい」というのは間違っていない。
(東京大学教養学部統計学教室編「基礎統計学 I 統計学入門」125 ページの記述によると、実際に過去の「事故間隔」を横軸、その頻度を縦軸にグラフにプロットしてみると、確かに指数分布が当てはまっているそうである)
2. 正しくない。その月に一度使ったコードを月内には2度と使わないことにすると、その時点で、そのコードが選ばれる確率は0となり、他の各コードが選ばれる確率が少しずつ増えることになるので、暗号を見破るヒントを与えてしまう。無作為抽出にするためには、すでに一度使ったコードも、次の抽出の機会には平等なチャンスで選ばれるようにしなければならない。
(ドイツ軍は、実際にこの誤りを犯し、そのためにイギリス軍に暗号を見破りやすくさせてしまった、といわれている)

2.

1. 年収の少ない人は、平均年収を書かないと思われるし、人は見栄をはるから正確なことを書かないおそれも高い。そもそも、同窓会の返信はがきで、年収のようなプライベートなことを尋ねてはいけない。
2. このことが言えるには、「国会議員になる資質がある人たち」という母集団が存在し、現実の国会議員が、その母集団から無作為抽出された標本になっていなければならない。しかし、「国会議員になる資質がある人たち」という母集団はおそらく存在しない。
3. 家庭用の照明ならば、切れてから取り替えればよいのだから、平均寿命が長い電球のほうが、長期的に見れば家計の節約になる。しかし、信号機の電球などの「絶対に故障してはいけない機器」では、故障しないうちに一定の間隔で電球を交換することになるから、絶対に故障しないことが保証される期間、すなわち「最低保証寿命」が長い電球のほうが、平均寿命が長い電球よりも望ましい。
4. 無事帰還した乗務員は攻撃を免れたのであり、攻撃により損害をうけた爆撃機の大半は帰還しなかったであろう。だから、帰還した乗務員だけから攻撃について聴取しても、適切な対応はとれないと思われる（これは実話です）。
5. 母平均は、人が知らないだけで、実際にはひとつに決まっている。だから、「20以上30以下」という具体的区間を求めた段階で、母平均が「20以上30以下」であるかどうかはすでに決まっている。「95%信頼区間」とは、この方法で信頼区間を何度も求めると、そのうち95%は母平均を本当に含んでいる区間である、という意味である。

- 不正が発見されたのは「全品目のうち3%の品目」についてであり、不足の量がいくらだったかは調査されていない。もし、残りの97%の品目の内容量が表示よりもずっと多いのなら、消費者は平均して損などしていない、ということになる。