Statistika

kontrolní otázky

Mirko Navara

1A. Vysvětlete rozdíly mezi následujícími pojmy:

1. střední hodnota,
2. výběrový průměr,
3. realizace výběrového průměru.

1B. Jaký je rozdíl mezi nejvíce eficientním odhadem a nejlepším nestranným odhadem?

1A. Vysvětlete rozdíly mezi následujícími pojmy:

1. střední hodnota,
2. výběrový průměr,
3. realizace výběrového průměru.

1A. Vysvětlete rozdíly mezi následujícími pojmy:
1. střední hodnota,
2. výběrový průměr,
3. realizace výběrového průměru.

Střední hodnota nemusí existovat. Pokud existuje, nemusí jít určit; projevuje se pouze zprostředkovaně ve výsledcích náhodných pokusů.
Výběrový průměr je náhodná veličina vypočtená z náhodného výběru; vždy existuje.
Realizace výběrového průměru je číslo vypočtené z realizace náhodného výběru, sloužící k (realizaci) odhadu neznámé střední hodnoty.

1A. Vysvětlete rozdíly mezi následujícími pojmy:
1. střední hodnota,
2. výběrový průměr,
3. realizace výběrového průměru.

Střední hodnota nemusí existovat. Pokud existuje, nemusí jít určit; projevuje se pouze zprostředkovaně ve výsledcích náhodných pokusů.
Výběrový průměr je náhodná veličina vypočtená z náhodného výběru; vždy existuje.
Realizace výběrového průměru je číslo vypočtené z realizace náhodného výběru, sloužící k (realizaci) odhadu neznámé střední hodnoty.

1B. Jaký je rozdíl mezi nejvíce eficientním odhadem a nejlepším nestranným odhadem?

1A. Vysvětlete rozdíly mezi následujícími pojmy:
1. střední hodnota,
2. výběrový průměr,
3. realizace výběrového průměru.

Střední hodnota nemusí existovat. Pokud existuje, nemusí jít určit; projevuje se pouze zprostředkovaně ve výsledcích náhodných pokusů.
Výběrový průměr je náhodná veličina vypočtená z náhodného výběru; vždy existuje.
Realizace výběrového průměru je číslo vypočtené z realizace náhodného výběru, sloužící k (realizaci) odhadu neznámé střední hodnoty.

1B. Jaký je rozdíl mezi nejvíce eficientním odhadem a nejlepším nestranným odhadem?

Nejlepší nestranný odhad je nejvíce eficientní mezi *nestrannými*, ale nemusí být nejvíce eficientní mezi všemi.

2A. Kdy nelze použít centrální limitní větu?

2B. Co lze říci o poměru rozptylu a střední hodnoty chí-kvadrát rozdělení?

2A. Kdy nelze použít centrální limitní větu?

2A. Kdy nelze použít centrální limitní větu?

Pro ta rozdělení, která nelze normovat, tj. pokud nemají rozptyl (speciálně, pokud nemají střední hodnotu), nebo je nulový (Diracovo rozdělení).

Dále ji nelze využít, pokud počet opakování pokusu je omezen natolik, že rozdělení (normovaného) výběrového průměru se od asymptotického podstatně liší; pak sice věta platí, ale nám nepomůže.

2A. Kdy nelze použít centrální limitní větu?

Pro ta rozdělení, která nelze normovat, tj. pokud nemají rozptyl (speciálně, pokud nemají střední hodnotu), nebo je nulový (Diracovo rozdělení).

Dále ji nelze využít, pokud počet opakování pokusu je omezen natolik, že rozdělení (normovaného) výběrového průměru se od asymptotického podstatně liší; pak sice věta platí, ale nám nepomůže.

2B. Co lze říci o poměru rozptylu a střední hodnoty chí-kvadrát rozdělení?

2A. Kdy nelze použít centrální limitní větu?

Pro ta rozdělení, která nelze normovat, tj. pokud nemají rozptyl (speciálně, pokud nemají střední hodnotu), nebo je nulový (Diracovo rozdělení).

Dále ji nelze využít, pokud počet opakování pokusu je omezen natolik, že rozdělení (normovaného) výběrového průměru se od asymptotického podstatně liší; pak sice věta platí, ale nám nepomůže.

2B. Co lze říci o poměru rozptylu a střední hodnoty chí-kvadrát rozdělení?

Je vždy 2; lze to říci, neboť chí-kvadrát rozdělení je z definice bezrozměrné.

3A. Proč máme nestranný odhad rozptylu a nemáme nestranný odhad směrodatné odchylky?

3B. Co potřebujeme k odhadu kvantilu pomocí kvantilu empirického rozdělení?

3A. Proč máme nestranný odhad rozptylu a nemáme nestranný odhad směrodatné odchylky?

3A. Proč máme nestranný odhad rozptylu a nemáme nestranný odhad směrodatné odchylky?

Měli jsme to štěstí, že rozptyl empirického rozdělení byl sice vychýlený, ale jeho poměr ke správné hodnotě závisel jen na rozsahu výběru, a ten známe.
Směrodatná odchylka empirického rozdělení závisí i na dalších okolnostech, které nelze obecně kompenzovat.

3A. Proč máme nestranný odhad rozptylu a nemáme nestranný odhad směrodatné odchylky?

Měli jsme to štěstí, že rozptyl empirického rozdělení byl sice vychýlený, ale jeho poměr ke správné hodnotě závisel jen na rozsahu výběru, a ten známe.
Směrodatná odchylka empirického rozdělení závisí i na dalších okolnostech, které nelze obecně kompenzovat.

3B. Co potřebujeme k odhadu kvantilu pomocí kvantilu empirického rozdělení?

3A. Proč máme nestranný odhad rozptylu a nemáme nestranný odhad směrodatné odchylky?

Měli jsme to štěstí, že rozptyl empirického rozdělení byl sice vychýlený, ale jeho poměr ke správné hodnotě závisel jen na rozsahu výběru, a ten známe.
Směrodatná odchylka empirického rozdělení závisí i na dalších okolnostech, které nelze obecně kompenzovat.

3B. Co potřebujeme k odhadu kvantilu pomocí kvantilu empirického rozdělení?

„Dostatečně velký rozsah výběru," což zde znamená, že v něm musí být „dost" výsledků větších i menších než odhadovaný kvantil.
„Dostatečný rozsah výběru" závisí na tom, jaký kvantil chceme odhadnout.

4A. Může existovat intervalový odhad náhodné veličiny, která nemá rozptyl, nebo dokonce střední hodnotu?

4A. Může existovat intervalový odhad náhodné veličiny, která nemá rozptyl, nebo dokonce střední hodnotu?

Může. Kvantily existují. Jen se musí odhadovat jinak.

5A. Čím to je, že pro horní odhad rozptylu používáme dolní kvantil?

5B. Náhodná veličina je čas. V jakých jednotkách se mohou měřit veličiny, používané pro odhad parametrů

1. metodou momentů,
2. metodou maximální věrohodnosti pro diskrétní rozdělení,
3. metodou maximální věrohodnosti pro spojité rozdělení?

5A. Čím to je, že pro horní odhad rozptylu používáme dolní kvantil?

5A. Čím to je, že pro horní odhad rozptylu používáme dolní kvantil?

Protože použité rozdělení χ^2 vyjadřuje rozdělení výrazu, ve kterém je rozptyl v převrácené hodnotě.

5B. Náhodná veličina je čas. V jakých jednotkách se mohou měřit veličiny, používané pro odhad parametrů

1. metodou momentů, [s^k pro k-tý moment]
2. metodou maximální věrohodnosti pro diskrétní rozdělení, [1]
3. metodou maximální věrohodnosti pro spojité rozdělení? [s^{-n}]

6A. Kterou z následujících úloh byste řešili metodou momentů (MM), resp. metodou maximální věrohodnosti (MMV)?

1. Odhad pravděpodobností výsledků hodu nepravidelnou kostkou.

2. Odhad rozdělení dešťových srážek za 24 hodin.

3. Z výšky členů sportovního oddílu máme odhadnout sport, kterému se věnují.

6A. Kterou z následujících úloh byste řešili metodou momentů (MM), resp. metodou maximální věrohodnosti (MMV)?

1. Odhad pravděpodobností výsledků hodu nepravidelnou kostkou.
[MMV i MM se stejným výsledkem; jde o diskrétní rozdělení ničím neomezené.]
2. Odhad rozdělení dešťových srážek za 24 hodin.
[MM; jedná se o smíšené rozdělení.]
3. Z výšky členů sportovního oddílu máme odhadnout sport, kterému se věnují.
[MMV; vybíráme z konečně mnoha možností.]

7A. Co je příčinou nesymetrie, kdy nulovou hypotézu můžeme zamítnout, ale alternativní nikoli?

7B. Pokud nám ROC křivka nevychází dostatečně „dobrá“ (neprochází „blízko“ bodu $(0, 1)$, co máme dělat?

7A. Co je příčinou nesymetrie, kdy nulovou hypotézu můžeme zamítnout, ale alternativní nikoli?

7A. Co je příčinou nesymetrie, kdy nulovou hypotézu můžeme zamítnout, ale alternativní nikoli?

Kritérium pro rozhodování vychází z předpokladu platnosti *nulové* hypotézy a používá pouze pravděpodobnost chyby 1. druhu.
Nepracuje s předpokladem platnosti *alternativní* hypotézy a pravděpodobnost chyby 2. druhu nebere v úvahu.

7A. Co je příčinou nesymetrie, kdy nulovou hypotézu můžeme zamítnout, ale alternativní nikoli?

Kritérium pro rozhodování vychází z předpokladu platnosti *nulové* hypotézy a používá pouze pravděpodobnost chyby 1. druhu.
Nepracuje s předpokladem platnosti *alternativní* hypotézy a pravděpodobnost chyby 2. druhu nebere v úvahu.

7B. Pokud nám ROC křivka nevychází dostatečně „dobrá“ (neprochází „blízko“ bodu $(0, 1)$, co máme dělat?

7A. Co je příčinou nesymetrie, kdy nulovou hypotézu můžeme zamítnout, ale alternativní nikoli?

Kritérium pro rozhodování vychází z předpokladu platnosti *nulové* hypotézy a používá pouze pravděpodobnost chyby 1. druhu.
Nepracuje s předpokladem platnosti *alternativní* hypotézy a pravděpodobnost chyby 2. druhu nebere v úvahu.

7B. Pokud nám ROC křivka nevychází dostatečně „dobrá“ (neprochází „blízko“ bodu $(0, 1)$, co máme dělat?

Zvětšit rozsah výběru nebo zvolit jiný test (popř. metodiku získávání dat, např. počkat, až budou rozdíly tříd výraznější).

8A. Kdy máme a kdy nemáme použít párový test?

8A. Kdy máme a kdy nemáme použít párový test?

Máme ho použít, pokud data jsou pořízena po dvojicích, na které působí stejný rušivý vliv, jehož vliv chceme eliminovat.

Nedá se použít, pokud data nejsou takto spárována (a už vůbec ne, pokud rozsahy souborů jsou různé).
Není vhodné ho použít ani tehdy, pokud data jsou pořízena po dvojicích, ale nepůsobí další rušivý vliv; snížili bychom sílu testu, neboť bychom použili $2\times$ méně dat.