

- 1) **Definujte pojem statistika.**
  - věda o sběru dat a zpracování hromadných údajů, zabývá se jevy, které mají hromadný charakter
  - hromadnost studována na statistických souborech
- 2) **Co je to popisná statistika?**
  - elementární metody sběru a zpracování informací
  - jednotkou je statistický soubor (osob, podniků, institucí, zvířat, zemí, atd.).
  - statistické soubory jsou tvořeny statistickými jednotkami, mají vlastnosti jednoznačně vymezeny.
- 3) **Co je matematická statistika a jak se dělí.**
  - moderní – zabývá se složitějšími metodami sběru a zpracování hromadných údajů; vytváří zvláštní druh matematických modelů – tzv. pravděpodobnostní modely – teorie pravděpodobnosti
- 4) **Typy statistických ukazatelů**
  - okamžikové, intervalové, primární, sekundární, extenzivní, intenzivní, stejnorodé, nestejnorodé
- 5) **Druhy statistických vlastností (2)**
- 6) **Statistické jednotky**
  - elementární jednotky stat. pozorování, jsou nositeli znaků
- 7) **Statistické znaky**
  - vlastnosti jednotek, která je předmětem zkoumání
  - kvalitativní – slovně vyjádřené – alternativní (2 obměny znaku); množné (více než 2 obměny)
  - kvantitativní – číselně vyjádřené, diskrétní (celočíslné), spojitě (desetinné číslo, logaritmy)
- 8) **Statistický soubor**
  - množina jedinců, na kterých je prováděno statistické šetření
  - základní soubor – všechny jednotky s danou vlastností
  - výběrový soubor – vybrán ze základního, podmnožina je menší
- 9) **Rozdíl mezi ZS a VS**
  - VS je vlastně část ZS
  - VS je menší než ZS (úplné zjišťování, tvořen všemi jednotkami), VS(neúplné zjišťování)
- 10) **Základní etapy statistických prací**
  - statistické šetření (zjišťování) - získávání neznámých informací o znacích statistických jednotek , výsledkem statistického zjišťování jsou neuspořádané údaje
  - statistické zpracování
  - statistická analýza
- 11) **Co je statistické zjišťování?**
  - získávání neznámých informací o znacích statistických jednotek
  - výsledkem statistického zjišťování jsou neuspořádané údaje
  - ankety, dotazníky, experiment, výsledek vědeckého experimentu
  - pro přehlednění se data třídí
- 12) **Základní míry polohy rozdělení a k čemu slouží**
  - průměry: aritmetický; vážený ar. prům.; harmonický; geometrický; celkový ar. prům.; chronologický
  - ostatní střední hodnoty: medián, modus
  - měly by jedním číslem popsat střední úroveň hodnoty statistického znaku a umožnit jeho hlubší analýzy
  - reprezentují vhodnou střední hodnotu daného souboru kolem níž se soustřeďují hodnoty tohoto souboru
- 13) **Prosté x vážené charakteristiky polohy – rozdíl (3)**
  - prosté – u neseřazených dat, máme-li relativní četnosti
  - vážené – u seřazených dat (tabulka rozdělení četností)
- 14) **Průměr aritmetický**
  - součet všech hodnot znaků dělený počtem znaků
- 15) **Průměr geometrický**
  - n-tá odmocnina ze součinu znaků
- 16) **Jaké znáte míry založené na geometrickém průměru?**
  - všude tam, kde má smysl násobit hodnoty, např. průměrný koeficient růstu nebo Fisherův index
- 17) **V jakém oboru statistiky se můžeme setkat s geometrickým průměrem**
  - používá se u časových řad – průměrné tempo růstu (koeficient růstu)
- 18) **Průměr harmonický**
  - podíl počtu pozorování a sumy převrácených hodnot znaků
- 19) **Kdy a k čemu používáme harmonický průměr (3)**
  - v indexní analýze; průměr převrácených hodnot
- 20) **Průměr chronologický**
  - použití v okamžikové časové řadě
  - prostá forma – tam, kde délka mezi rozhodnými obdobími je stejná)
  - vážená forma – kde vahami jsou počty dní v měsíci,...

- 21) Tempo a průměrný koeficient tempa růstu**  
- počítání geometrického průměru
- 22) Medián**  
-  $x$  s vlnkou - prostřední hodnota znaku v souboru uspořádaná podle velikosti  
- lichý počet hodnot v souboru - střední hodnota  
- sudý počet hodnot - průměr střední hodnoty
- 23) Modus**  
- hodnota, která se nejčastěji vyskytuje, hodnota znaku s největší četností
- 24) Jak vypočítáte modus a medián spojitě náhodné veličiny, znáte-li její distribuční funkci?**
- 25) Uveďte situaci, kdy může medián popsat polohu statistického souboru lépe než průměr.**  
- medián může popsat polohu statistického souboru lépe, pokud je nějaká hodnota hodně vychýlená, tzn., že se hodně liší od ostatních  
- pak je průměr zkreslený a medián je lepší měrou polohy statistického souboru. př.: 4, 5, 5, 5, 5, 7, 48
- 26) Pro která pravděpodobnostní rozdělení je jejich střední hodnota rovna mediánu a zároveň modu? Vysvětlete a uveďte příklady.**  
Pro symetrická (Normální, studentovo)
- 27) K čemu se používají podmíněné průměry**  
- je to nejjednodušší způsob určení regresní závislosti (přímka podmíněných průměrů)- nelze však na jejich základě provádět odhady
- 28) Co se stane s průměrem, rozptylem, směrodatnou odchylkou, mediánem a rozpětím statistického souboru, jestliže každá hodnota statistického souboru se:**  
**a) zvětší dvakrát** - průměr a medián se zdvojnásobí, rozptyl se zvýší čtyřikrát; směrodatná odchylka a rozpětí statistického souboru se zvýší dvakrát  
**b) zvětší o čtyři** - průměr a medián se zvětší o čtyři; rozptyl se nezmění; směrodatná odchylka a rozpětí statistického souboru se nezmění
- 29) Rozdělení četnosti a co je intervalové rozdělení četnosti**  
- rozdělení četnosti - u nespojitých znaků původně neuspořádané údaje roztrždit do rozdělení četnosti
- 30) Jak se stanovuje interval relativní četnosti ZS u malých VS**  
- výběr relativní četnosti se řídí binomickým rozdělením v případě výběru bez vracení se řídí hyperbolickým rozdělením  
- výpočet vede ke složitým variacím, proto máme sestaveny tabulky a přímo odečítáme meze intervalu z tabulek
- 31) Definice pojmu kumulativní četnost**  
- absolutní a relativní  
- vznikají postupným načítáním
- 32) Druhy grafů**  
- spojnicové, sloupcové (polygon, histogram), bodové, výšecové, speciální (kvartogram)
- 33) Histogram**  
- sloupcový graf - u intervalového rozdělení četností
- 34) Které charakteristiky statistického souboru můžete přibližně zjistit z histogramu četnosti, aniž byste prováděli výpočet?**  
- počet intervalů a jejich šířku, absolutní četnost intervalu a pokud jsou intervaly stejně dlouhé i modus
- 35) Jaký graf používáme u jednorozměrných četností.**  
- sloupcový
- 36) Základní míry variability**  
- absolutní: rozptyl, směrodatná odchylka, variační rozpětí, prům. odchylka  
- relativní: variační koeficient, relativní průměrná odchylka
- 37) Rozptyl**  
- aritmetický průměr čtverců individuálních odchylek jednotlivých hodnot znaku od aritmetických průměrů  
- nedostatek - jednotky jsou druhou mocninou původních jednotek
- 38) Směrodatná odchylka v souboru výběrových průměrů**  
- měří abs. Variabilitu  
- je uvedena ve stejných měrných jednotkách jako zkoumaný stat. znak;  $s = \text{odm.s na } 2$   
- prostá:  $S_0 = \text{odm. } z((\sum (x_i - \bar{x})^2) / n)$   
- vážená:  $S_0 = \text{odm. } z((\sum (x_i - \bar{x})^2 \cdot n_i) / (\sum n_i))$   
- informuje o proměnlivosti jednotlivých hodnot znaku kolem výběr. aritm. průměru
- 39) Variační rozpětí**  
- jednoduchá míra adaptability  
- pouze odchylky mezi sebou - orientační
- 40) Relativní ukazatele variability.**  
- variační koeficient, relativní průměrná odchylka
- 41) K čemu slouží variační koeficient? Jaká je jeho přednost?**  
- variační koeficient je základní mírou relativní variability

- může se použít i tehdy pokud se znaky liší svou úrovní, což je výhoda
  - počítá se jako podíl směrodatné odchylky a průměru
- 42) **Jak se změni variační koeficient, přičteme-li ke všem hodnotám souboru stejnou konstantu?**  
 Směrodatná odchylka v čitateli zůstane stejná a průměr ve jmenovateli se zvětší tuto konstantu => variační
- 43) **Lze vždy vypočítat variační koeficient souboru dat? Názor zdůvodněte.**  
 Ne. Variační koeficient se počítá jako podíl směr. odchylky a průměru => je-li např. průměr nulový, Variační koeficient vypočítat nelze.
- 44) **Kvantil**  
 - je hodnota, která rozděluje soubor hodnot na dvě části
- 45) **Kvartil**  
 - dělí soubor po 25%
- 46) **Rozdíl mezi charakteristikami šikmosti a špičatosti (3)**  
 - charakteristika šikmosti (nesouměrnosti) – ukazuje, jak soubor vypadá, stupeň koncentrace malých a velkých hodnot v souboru  
 - charakteristika špičatosti – ukazuje, jak jsou hodnoty nahloučeny kolem průměru
- 47) **Význam výběrového šetření v praxi (3)**  
 - pořizujeme výběrový soubor, aby nám poskytl informace o celém souboru  
 - hlavním nedostatkem je, že jsou zatíženy výběrovou chybou
- 48) **Výhody úplného zjišťování oproti neúplnému výběrovému zjišťování**  
 - úplné – při práci se základním souborem, nákladné, zdlouhavé, občas nemožné  
 - neúplné – při práci s výběrovým souborem, výběrový soubor musí být dobrým reprezentantem
- 49) **Vysvětlíte pojmy oblastní a vícestupňový náhodný výběr**  
 - vícestupňový - výběr provádíme na více stupních (města – školy – fakulty – ročníky – studenti)  
 - oblastní - dvoustupňový výběr; v 1. stupni vybíráme oblast a ve 2. stupni vybíráme z oblasti jednotku
- 50) **Kvótní výběr - v čem spočívá**  
 - typ mechanického výběru při náhodném výběru
- 51) **Jaké znáte techniky pořízení náhodného výběru?**  
 - losování – opora výběru – výběr zastoupíme lístky  
 - tabulky náhodných čísel – generátor náhodných čísel  
 - mechanický výběr – systematické, každá n-tá jednotka v náhodně uspořádané posloupnosti speciální výběr
- 52) **Existuje rozdíl mezi stanovením intervalu u vracení a bez vracení?**  
 - s vracením – jednotku po výběru vracíme zpět  
 - bez vracení – rozsah ZS se zmenšuje, pravděpodobnost vybrání se zvětšuje  
 - u velkých souborů zbytečné pracovat s vracením
- 53) **Jaký test k ověřování náhodnosti výběrového souboru? (3)**
- 54) **Metoda základního masivu**  
 - kdy se soubor skládá z několika velkých a mnoha malých jednotek  
 - zjišťování provádíme na velkých jednotkách
- 55) **Záměrný výběr**  
 - značná míra subjektivity toho, kdo vybírá  
 - vybere ty, o kterých si myslí, že dobře zastoupí soubor, ty blízké průměru, nelze vyvodit chyba
- 56) **Dělení (druhy) náhodného výběru**  
 - s vracením, bez vracení  
 - s nestejnou pravděpodobností vybrání  
 - prostý – se stejnými pravděpodobnostmi
- 57) **Náhodný jev**  
 - jev, který může nastat nebo nenastane v závislosti na náhodě a je výsledkem náhodného pokusu (charakterizuje výsledek náhodného pokusu kvalitativně)
- 58) **Náhodný pokus**  
 - realizace podmínek a vlivů, z nichž některé jsou známe a jiné náhodné
- 59) **Jev jistý, náhodný, nemožný**  
 - jev jistý - takový, který vždy nastane při každém provedení náhodného pokusu  
 - jev náhodný - jevy, které v závislosti na náhodě mohou, ale nemusí při uskutečňování daného komplexu podmínek nastat  
 - jev nemožný - náhodný jev, který nenastane při žádném provedení náhodného pokusu
- 60) **Klasické a statistické definice pravděpodobnosti**  
 - klasická - může li určitý pokus vykazat konečný počet n různých výsledků, které jsou stejně možné a jestliže m těchto výsledků má za následek nastoupení jevu A, kdežto zbylých n-m vylučuje: potom  $P(A)=m/n$   
 - statistická - spojena s pojmem relativní četnosti; s rostoucím počtem pokusů se relativní četnost stabilizuje a přibližuje se k určitému konstrukčnímu číslu.  $P(A)=\lim_{n \rightarrow \infty} \frac{m}{n}$
- 61) **Matematická charakteristika pravděpodobnosti**

**62) Rozdíl mezi náhodnou veličinou a náhodným jevem (2)**

- náhodný jev – takový jev, který v závislosti na náhodě může, ale nemusí při uskutečňování daného komplexu podmínek nastat; charakterizuje výsledek náhodného pokusu kvalitativně (slovně)
- náhodná veličina – libovolná kvantitativní charakteristika náhodného pokusu; proměnná, která nabývá konkrétních hodnot, či hodnot z různých intervalů v závislosti na náhodě

**63) Zákon rozdělení náhodné veličiny**

- pravidlo, které každé hodnotě, nebo množině hodnot z každého intervalu přiřazuje pravděpodobnost, že náhodná veličina nebude této hodnoty, nebo hodnoty z tohoto intervalu
- tento zákon může být vyjádřen různou formou:  
jako řada rozdělení pravděpodobností (grafem je polygon, diskrétní veličiny)  
distribuční fce (univerzální zákon rozdělení, diskrétní i náhodné veličiny)  
hustota pravděpodobnosti (spojité náhodné veličiny)

**64) Druhy rozdělení náhodných veličin**

- spojité (normální, exponenciální, chí-kvadratické, Studentovo t-rozdělení, F-rozdělení, rovnoměrné rozdělení)
- nespojité - diskrétní (Alternativní, Binomické, Poissonovo, Hypergeometrické, Geometrické)

**65) Charakterizujte normální rozdělení**

- náhodná veličina se řídí normálním rozdělením, je-li její střední hodnota  $\mu$  a rozptyl  $\sigma^2$
- grafem hustoty pravděpodobnosti je Gaussova křivka
- speciálním případem normálního rozdělení je normované normální rozdělení

**66) Binomické rozdělení je (možnosti) (2)**

- nejdůležitější typ rozdělení diskrétní náhodné veličiny
- rozdělením náhodné veličiny, která představuje počet výskytů jevu A při n nezávislých pokusech, přičemž pravděpodobnost výskytu jevu A je v každém pokusu konstantní

**67) Jaký je vztah binomického a Poissonova rozdělení?**

- má-li náhodná veličina X binomické rozdělení takové, že počet pokusů n je dostatečně veliké (nad 30), pravděpodobnost výskytu sledovaného jevu v jednom pokuse pod 0,1 a n  $\rightarrow$  konečné číslo, je možno toto rozdělení aproximovat Poissonovým rozdělením

**68) Pravidlo tří sigma**

- i když náhodná veličina X, která má normální rozdělení, může nabývat hodnot z intervalu od  $(-\infty, \infty)$ , je téměř nemožné, aby se pozorované hodnoty této veličiny odchylovaly od střední hodnoty o více než 3 sigma

**69) Co vyjadřuje zákon velkých čísel?**

- se zvyšováním počtu náhodných pokusů dochází k přibližování se empirické charakteristiky popisující výsledky těchto pokusů k charakteristice teoretické

**70) Co vyjadřuje centrální limitní věta?**

- vyjadřuje konvergenci pravděpodobnostních rozdělení k normálnímu rozdělení při dostatečně velkém rozsahu souboru.

**71) Co je normovaná náhodná veličina, jaké má charakteristiky a jaký má význam?**

- má normální rozdělení se střední hodnotou 0 a rozptylem 1
- význam je ve výpočtu distribuční funkce, která se z normálního rozdělení počítá obtížně

**72) V čem spočívá z pohledu teorie pravděpodobnosti průnik jevů A,B a sjednocení jevů A,B**

- průnik jevů A,B - spočívá v současné realizaci jak jevu A, tak jevu B
- sjednocení jevů A,B - spočívá v nastoupení alespoň jednoho z jevů A nebo B

**73) Je možné, aby existovaly 2 náhodné jevy, že pravděpodobnost jejich průniku je větší než pravděpodobnost jejich sjednocení?**

- Ne. Protože pravd. průniku může být max. rovná pravděp. sjednocení, když množiny splývají nebo je menší nebo množiny nemají průnik.

**74) Při výpočtu pravděpodobnosti projiti třemi zkouškami, z nichž každá má svou pravděpodobnost úspěchu používáme:**

- a) sčítání
- b) násobení
- c) rozdíl

**75) Podmínky pro sčítání pravděpodobností a vzorec**

- jsou-li jevy A a B slučitelné, potom pravděpodobnosti jejich sjednocení se rovná součtu pravděpodobností jednotlivých jevů zmenšenému o pravděpodobnost jejich průniků
- v případě neslučitelných jevů je průnik těchto jevů jev nemožný

**76) Co jsou kvalitativní znaky, jak se dělí, příklady**

- kvalitativní znaky jsou znaky slovní, získané z anket, dotazníků
- dělíme na znaky alternativní -

**77) Jak ověříte nezávislost dvou kvalitativních znaků?**

Kontingenční tabulkou

**78) Jakým způsobem můžete zjistit, jestli existuje závislost mezi dvěma kvalitativními znaky a jakým v případě kvantitativních znaků?**

- závislost mezi 2 kvalitativními znaky ověřujeme kontingenční tabulkou a testem chí-kvadrát.
- závislost mezi 2 kvantitativními znaky měříme klasicky pomocí regresní a korelační analýzy, celkový F-test, Test t pro

jednotlivé parametry a koeficienty

- u 1 kvalitativního a 1 kvantitativního znaku se používá jednofaktorová analýza rozptylu

**79) Koeficient asociace slouží k:**

- vyjádření těsnosti 2 alternativních znaků

**80) Jak určíme nejvhodnější typ funkce při měření závislosti dvou kvantitativních znaků**

- zkušenost, logika, emp. metoda-korelační pole

- zkoušet = počítat - zpětně vybrat ten s nejvyšším korelačním charakterem

**81) Jaké jsou hlavní úlohy při měření závislosti 2 kvantitativních znaků**

- vystihnout průběh závislosti závisle proměnné na nezávisle proměnné, tak abychom mohli provádět odhady závisle proměnné na základě daných hodnot nezávisle proměnné

- změřit sílu závislosti, abychom mohli posoudit její sílu, intenzitu a abychom mohli zároveň posoudit přesnost odhadů z 1 bodu

- 1. úkol - regrese, 2. úkol - korelace

**82) Teoretický soubor výběrových průměrů**

- ze základního souboru vybereme všechny teoreticky možné VS, těch je nekonečně mnoho; v každém výběrovém souboru si vypočítáme výběrový průměr, všechny tyto průměry nám vytvoří teoretický soubor výběrových průměrů

**83) Statistická indukce**

- nejprve pořídíme výběrový soubor, na základě VS si spočítáme výběrové charakteristiky, na základě výběrových charakteristik odhadujeme charakteristiky ZS

**84) Jaké známe odhady (3)**

- bodový – jedno konkrétní číslo, které vybereme z VS, aby nám nahradilo ZS

- intervalový – stanovení intervalu, ve kterém ta neznámá charakteristika bude ležet, a určitou pravděpodobností

**85) Co je bodový odhad?**

- bodový – jedno konkrétní číslo, které vybereme z VS, aby nám nahradilo ZS

**86) Jaké znáte vlastnosti bodových odhadů a co vyjadřují, jaké jsou na ně kladeny požadavky?**

- nezkreslenost, nestrannost odhadu (střední hodnota výběrové statistiky = odhadované charakteristice)

- konzistence (odhad se s rostoucím rozsahem výběru blíží odhadované charakteristice základního souboru)

- vydatnost (co nejmenší rozptyl)

- postačujícnost (mimo ní neexistuje žádná jiná statistika poskytující další doplňující informace o odhadované charakteristice základního souboru)

**87) Proč musí být bodový odhad v základním souboru vydatný**

- můžeme použít více charakteristik odhadu, za nejvydatnější je ta, která má nejmenší rozptyl

**88) Jak spolu souvisí přesnost odhadu a spolehlivost odhadu?**

- čím širší interval spolehlivosti, tím ale menší přesnost

- spolehlivost = pravděpodobnost, se kterou bude odhadovaná charakteristika ležet v tom vymezeném intervalu; = maximální chyba, které se při odhadu s danou spolehlivostí můžeme dopustit

**89) Co znamená, že odhad je vychýlený? Znáte některé vychýlené odhady?**

-  $E(g)$  -  $G$  je tzv. zkreslení neboli vychýlení

- takový odhad vede k systematickému nadhodnocování či podhodnocování odhadované charakteristiky ZS

**90) Přípustná chyba u intervalového odhadu a k čemu jí používáme**

- chyba, které se při odhadu můžeme dopustit, aby hodnota padla do intervalu

- přesnost intervalového odhadu je charakterizována přípustnou chybou odhadu delta, která představuje polovinu délky intervalu spolehlivosti

**91) Přesnost odhadu: (2)**

- pravděpodobnost, s jakou se charakteristika nachází v intervalu

- vyneseme kritickou hodnotu příslušného rozdělení

- max chyba, které se při odhadu s danou spolehlivostí dopustíme vyjádřena hodnotou směrodatné odchylky souboru výběrových prům.

- ani jedna správně

**92) Jaké znáte metody pro získání odhadů parametrů regresních funkcí lineárních v parametrech? Napište princip metod.**

- požadavek kompenzace kladné a záporné odchylky empirických hodnot od hodnot vyrovnaných a metoda nejmenších čtverců (aby součet čtverců popsaných odchylek byl minimální)

**93) Jaké znáte metody pro získání odhadů parametrů regresních funkcí nelineárních v parametrech?**

- pro funkce nelineární v parametrech používáme linearizující transformaci

- pak použijeme metodu nejmenších čtverců, parciální derivace, dále dostaneme soustavu normálních rovnic a nakonec pomocí Cramerova pravidla (determinanty) vyjádříme  $b_0$ ,  $b_1$ , ...

**94) U kterých z uvedených regresních funkcí lze k odhadu parametrů použít metodu nejmenších čtverců: přímka, parabola, exponenciála? Názor vysvětlete.**

- u přímky a paraboly, protože jsou lineární v parametrech na rozdíl od exponenciály, která není - tam je nutná lineární transformace.

**95) Pojmy: (3)**

a) alternativní hypotéza – popírá platnost nulové hypotézy

- b) testovací kritérium – míra nesouhlasu výsledků pokusu s testovanou hypotézou (odpovídají-li data nulové hypotéze- testovací kritérium je rovno nule; čím více se výběrové hodnoty blíží k alternativní hypotéze, tím roste i testovací kritérium)
- c) hladina významnosti – pravděpodobnost chyby 1. druhu; udává výši rizika, s jakým se  $H_0$  zamítá, i když platí
- 96) Alternativní hypotéza**
- popírá platnost nulové hypotézy
  - přijímáme ji jestliže jsme nulovou hypotézu zamítli jako nesprávnou
- 97) Vysvětlete pojem hladina významnosti a kritický obor.**
- hladina významnosti - pevná, předem zvolená pravděpodobnost chyby I. Druhu
  - kritický obor - podprostor výběrového prostoru obsahující hodnoty svědčící ve prospěch alternativní hypotézy
- 98) Vysvětlete pojem "síla testu" a "hladina významnosti".**
- síla testu - s jakou pravděpodobností se nedopustíme chyby II. druhu
  - hladina významnosti - pevná, předem zvolená pravděpodobnost chyby I. druhu
- 99) Co udává hladina významnosti alfa**
- udává výši rizika s jakým  $H_0$  zamítá, i když platí pravděpodobnost chyby 1. druhu;  $\alpha = P(T \text{ je elem. } K \setminus H_0)$
- 100) Testovací kritérium je**
- veličina vypočtená z výběrových hodnot, řídí se rozdělením (norm. studentovým), porovnáváme s krit. hodnotou
- 101) Interval spolehlivosti a k čemu slouží**
- interval, ve kterém leží neznámé charakteristiky s určitou předem známou pravděpodobností. Prostřednictvím intervalu spol. posuzujeme přesnost odhadu. s rostoucí šířkou intervalu spol. klesá přesnost odhadu.
- 102) Čím můžeme ovlivnit velikost intervalu spolehlivosti?**
- zvolenou hladinou spolehlivosti. Čím širší interval spolehlivosti, tím ale menší přesnost.
- 103) Vysvětlete pojmy chyba I. druhu, chyba II. druhu a síla testu.**
- Zamítnutí nulové hypotézy, ačkoliv ve skutečnosti platí
  - přijmutí nulové hypotézy, ačkoliv ve skutečnosti platí hypotéza alternativní
  - s jakou pravděpodobností se nedopustíme chyby II. druhu
- 104) Jakým způsobem můžeme snížit pravděpodobnost chyby I. druhu a jak pravděpodobnost chyby II. druhu (při dané chybě I. druhu)?**
- chybu I. druhu snížíme tak, že místo 95% hladiny významnosti volíme 99%
  - chybu II. druhu snížíme tím, že zvýšíme rozsah souboru
- 105) Rozdíl mezi dvoustranným a jednostranným intervalem**
- dvoustranný - charakteristiky ZS jsou omezeny zdola i shora
  - jednostranný - charakteristiky ZS jsou omezeny shora nebo zdola
- 106) Obecný postup statistického testování**
- vybereme vhodný test (parametrický a neparametrický)
  - formulace nulové a alternativní hypotézy
  - volba hladiny významnosti
  - volba testovacího kritéria
  - určení kritického oboru
  - výpočet hodnoty testovacího kritéria z výběrových hodnot
  - rozhodnutí: jestliže vyp. hodnota t.k. padne do kritického oboru  $H_0$  se zamítá jinak se  $H_0$  nezamítá
- 107) Rozdíl mezi parametrickými a neparametrickými testy**
- Parametrické - testy sloužící k ověřování hypotéz, které se týkají hodnot parametrů rozdělení - spolehlivé, vyžadují znalost ZS a hodnotu parametrů
  - Neparametrické - místo původních hodnot pracujeme s pořadovými čísly - jednoduché, menší síla testu, nevyžadují znalost ZS
- 108) Jaký závěr lze učinit, vede-li celkový F-test o regresní funkci k zamítnutí testované hypotézy (nejdříve uveďte, jak jsou při tomto testu formulovány testovaná a alternativní hypotéza).**
- testovaná = neexistence vlivu faktoru (o nezávislosti znaku Y na zkoumaném faktoru); Alternativní = existence
  - závěr = podobnost skupinových průměrů na celý soubor
- 109) Lze zvětšením rozsahu výběru ovlivnit šíři intervalu spolehlivosti odhadu střední hodnoty? Pokud ano, vysvětlete jak.**
- pokud máme větší rozsah souboru, vede to k užšímu intervalu spolehlivosti. Vzorek je více interpretativní.
- 110) "Jestliže hypotézu  $\mu = \mu_0$  zamítneme při jednostranném testu na určité hladině významnosti, pak ji zcela určitě zamítneme i při oboustranné alternativě při stejné hladině významnosti." Souhlasíte? Vysvětlete proč.**
- neplatí vždy  $\leq$  dvoustranný test má při stejné hladině významnosti poloviční  $\alpha$  a hodnota se do něj může, ale nemusí vejít
- 111) Postup při použití testu minimální průkazové definice**
- u vyváženého modelu: dělíme difference mezi průměry ale neseřazujeme je podle velikosti; spočítáme hodnotu  $D_{\min} = g = \text{odmocnina } z \left( \frac{sr}{n/2} \right)$ ; porovnáme difference s touto hodnotou; jestli je dif. větší než  $D_{\min}$  tak je rozdíl mezi průměrem statisticky významný

**112) Při testování spotřeby automobilu určité značky používáme:**

- a) F-test
- b) t-test
- c) Bartlettův test

**113) Při testování více než dvou rozptylů se nepoužívá**

- a) Hartleyův test
- b) Bartlettův test
- c) F-test

**114) Bartelův test**

- test rozptylu více souborů,  $H_0$  rozptyly se rovnají
- testuje, zda můžeme udělat analýzu rozptylů - podmínka - shodné rozptyly z normálního rozdělení
- pokud nesplňuje: a) transformace hodnot; b) doměření hodnot; vyloučení ext. hodnot

**115) Cochranův test**

- stejné rozsahy n

**116) Hartleyův test**

- řídí se f rozdělením

**117) Ferrarův – Glauberův test, k čemu se používá**

**118) Rozdílové testy**

**119) Pořadové testy**

**120) Rozptyl a jak jej definujeme**

- aritmetický průměr čtverců individuálních odchylek jednotlivých hodnot znaku x od aritmetických průměrů
- charakteristika variability
- informuje o proměnlivosti jednotlivých hodnot znaku kolem výběr. ar. prům.
- výběrový rozptyl s na 2 souboru: pozorované  $x_1, x_2, \dots, x_n$

**121) Jak provádíme rozklad celkového rozptylu? (4)**

- rozklad celkového rozptylu všech hodnot na dílčí rozptyly (vyjadřují podíl jednotlivých faktorů na celkovém rozptylu) a na rozptyl reziduální (zbytkový, vyjadřuje podíl faktorů působících náhodně)

**122) Kde prakticky využíváme rozklad rozptylu?**

- při počítání vnitro a meziskup. variability
- u regres. a korel. analýzy ke konstrukci indexů

**123) K čemu slouží analýza rozptylu a jaké jsou předpoklady jejího použití**

- zobecnění t-testů na více než 2 VS rozklad výběr. rozptylu na něk. částí, které jsou příslušné jednotlivým uvažovaným zdrojům variability, zkoumáme li vliv jednoho či více faktorů na výsledky kvantitativního znaku
- analýza rozptylu je založena na předpokladu, že ze zjištěných hodnot lze provést odhad rozptylů - 2 způsoby: uvnitř tříd, mezi třídami
- typické použití v zemědělství - plemenictví

**124) Analýza rozptylu je (možnosti) (2)**

- je obecný postup, který patří do oblasti statistických metod – metodologický nástroj
- technika pro zakládání a vyhodnocování experimentu
- je více výběrový test, který testuje rozdíl mezi průměry
- zkoumá vliv jednoho i více faktorů na výsledný znak kvantitativní, každý faktor je sledován na několika úrovních (třídách), každá úroveň představuje 1 výběrový soubor; faktorem může být znak kvantitativní i kvalitativní
- pokud je úroveň každého faktoru pevně fixována – model analýzy rozptylu s pevnými efekty
- pokud jsou úrovně faktorů náhodně vybrány – model s náhodnými efekty

**125) Jaké míry jsou založeny na rozkladu rozptylu?**

- korelační koeficienty

**126) Jakou metodu byste použili k posouzení, jestli je prodej určitého zboží v průběhu dne rovnoměrně rozložen?**

Možná rozptyl - nebo řetězový index

**127) Kde prakticky využíváme rozklad rozptylu?**

- při počítání vnitro- a meziskup. variability a u regres. a korel. analýzy ke konstrukci indexů.

**128) Metody podrobnějšího hodnocení výsledků analýzy rozptylu**

- metody mnohonásobných porovnání, s-metoda (Scheffeho metoda), t-metoda (metoda minimální důkazové difference), Cramerova metoda, Duncanova metoda

**129) Duncanova metoda**

**130) Postup při použití Duncanova testu**

- hodnotám přiřadíme pořadové číslo a spočítáme difference
- uspořádáme do tabulky, určíme krit. hodnotu difference a porovnáme difference s kritickými, jeli dif. větší než krit. pak je rozdíl statisticky významný

**131) Kramerova metoda a k čemu slouží**

- pomocí ní se provádí podrobnější hodnocení výsledku v analýze rozptylu u neváženého modelu
- rozdíl mezi průměry je statisticky významný, jestliže jejich difference je větší než výraz na pravé straně nerovnosti

**132) Scheffeho metoda**

- rozdíl mezi průměry je stat. významný, jestliže jejich difference je větší než vypočítaná hodnota

**133) Tukeyův test**

**134) Rozdíl mezi vyváženým a nevyváženým modelem analýzy rozptylu**

- vyvážený - stejný počet opakování v každém řádku
- nevyvážený - různý počet opakování v ...

**135) Rozdělení F při analýze rozptylu**

- $F_{\alpha}(m-1; n-m)$
- rozdělení F pro  $f_1=m-1$  a  $f_2=n-m$  stupňů volnosti

**136) K čemu se používá F-test**

- test o rozdílu 2 výběrových rozptylů

**137) Jaké testy předcházejí analýze rozptylu**

- Bartlettův
- Cochranův
- Hartleyův

**138) Kdy používáme analýzu rozptylu + podmínky použitelnosti**

- používá se, když zkoumáme vliv jednoho či více faktorů na výsledný znak kvantitativní
- podmínky použitelnosti - normalita rozdělení ZS z kt. jsou porovnány VS; statistická nezávislost náhodných chyb; existence shodných rozptylů ZS ze kt. jsou porovnány VS

**139) Co je interakce u AR?**

**140) Jak se provádí a proč podrobnější vyhodnocení AR?**

**141) Modely AR dvojného třídění**

- zobecnění t-testů na více než 2 VS rozklad výběr. rozptylu na něk. částí, které jsou příslušné jednotlivým uvažovaným zdrojům variability, zkoumáme li vliv jednoho či více faktorů na výsledky kvantitativního znaku
- analýza rozptylu je založena na předpokladu, že ze zjištěných hodnot lze provést odhad rozptylů - 2 způsoby: uvnitř tříd, mezi třídami
- typické použití v zemědělství - plemenictví

**142) Vypište neparametrické testy**

- testy dobré shody –  $\chi^2$  test, Kalgor-Smirnovův test
- vlastní neparametrické testy – Wilcoxon-Whiteův test (parametrický t-test), znaménkový test (parametrický pár. t-test), Wilcoxonův test (parametrický t-test), Kruskal-Wallisův test (parametrická analýza rozptylu), Dixonův test, Friedmanův test, test iterací

**143) Výhody a nevýhody neparametrických testů (3)**

- v praxi se často setkáváme s výběry poměrně malých rozsahů, které pocházejí z výrazně nenormálních souborů, nebo ze souborů, o jejichž rozdělení nic nevíme
- nepředpokládáme specifikované rozdělení základního souboru, z něhož se získává náhodný výběr
- výhoda – nezávislost na tvaru rozdělení studovaných náhodných veličin, použitelnost pro studium jak znaků kvantitativních, tak znaků kvalitativních (obecnější použití); výpočetní jednoduchost
- nevýhody – menší síla (schopnost odhalit nesprávnost testované hypotézy)

**144) Které testy používáme jako neparametrickou obdobu párového t-testu a jak je provádíme**

- znaménkový: spočítáme difference mezi páry hodnot; počet kladných hodnot  $z^+$ ; počet záporných hodnot  $z^-$ ; menší z obou je testovací kritérium  $z$ :  $z$  je větší než  $z_{\alpha}$  (sečtou se počty ve skupinách a min. výsledná hodnota je testové kritérium, které se porovná s tabulkovou hodnotou)
- wilcoxonův: spočítáme difference; k nenulovým diferencím přiřadíme vzestupně pořadová čísla; pořadová čísla rozřídíme zpětně do dvou skupin podle znamének diferencí; v obou skupinách pořadová čísla sečteme a dostaneme  $W^+$  a  $W^-$ ;  $W_{\min}(W^+, W^-)$  menší z obou čísel je test. krit.;  $W$  je menší než  $W_{\alpha}$  -  $H_0$  se zamítá;

**145) Který test používáme jako neparametrickou obdobu jednoduché analýzy rozptylu a jak se tento test provádí**

- Kruskal-Wallisův test: hodnoty sloučíme do jednoho výběru; seřadíme od největší do nejmenší; očíslovíme vzestupně; sečteme pořadová čísla pro každý řádek zvlášť; výpočet test. kritérium, porovnáme s kritickou hodnotou

**146) Postup při Wilcoxon-Whiteova testu**

- hodnoty sloučíme do jednoho výběru
- seřadíme od nejmenší do největší
- očíslovíme hodnoty vzestupně
- zpětně rozřídíme do VS
- sečteme pořadová čísla;  $T = \min(T_x, T_y)$ ;  $T$  je menší než  $T_{\alpha}$  -  $H_0$  se zamítá
- u souboru  $n$  větší než 20 normální aproximace odhadnem



- 147) Test o shodě průměru základního souboru (2)**
- je určen k testování rozdílů mezi průměry u nezávislých výběrů
  - používá se v případě párových výběrů
  - používá se k testování rozdílu mezi výběrovým prům. a předpokládanou hodnotou prům. základního souboru
  - určuje, zda prům. zákl. souboru je stat. významný
  - ani jedna správně
- 148) K čemu je Friedmanův test? (4)**
- máme více než 2 závislé soubory, tzn.  $k > 2$
- 149) Uveďte aspoň 3 testy, u nichž použijete jako kritické hodnoty kvantily rozdělení chí-kvadrát. Uveďte, čeho se týkají a co říkají.**
- chí-kvadrát dobré shody
  - test o parametru delta exponenciálního rozdělení
  - test hypotézy o rozptylu v základním souboru
- 150) Jaké znáte testy shody? K čemu slouží?**
- Chí-kvadrát test dobré shody, shoda 2 průměrů, rozptylů
  - porovnávají dva různé soubory
- 151) Uveďte alespoň dva případy, kdy použijete při testování hypotézy testované kritérium, které má rozdělení chí-kvadrát.**
- chí-kvadrát dobré shody, test o parametru delta expon. rozdělení a test hypotézy o rozptylu v zákl. souboru.
- 152) Kdy používáme Fisherův test**
- když je  $n$  menší než 20 nebo  $n$  je prvkem (20,40) a jedina rel. čet. 15
- 153) V kterých tabulkách se provádí  $\chi^2$  test?**
- 154) Jaké jsou základní předpoklady Kolmogorov-Smirnovova testu**
- 155) Kdy používáme chí-kvadrát test?**
- 156) Jak dělíme metody vícerozměrné stat. analýzy z hlediska klasifikace jednotek**
- shluková analýza
  - diskriminační analýza
  - faktorová analýza...
- 157) Shluková analýza**
- vícerozměrná statistika
  - souhrnný název pro řadu výpočetních postupů, jejichž cílem je rozklad daného souboru na několik relativně homogenních množin (shluků) a to tak, aby jednotky uvnitř jednotlivých shluků si byly co nejvíce podobné a jednotky patřící do jiných shluků co nejvíce nepodobné
  - míra vzdálenosti - je jejím určením
  - vytváří celky metody - nejbližšího souseda, nejvzdálenějšího souseda, průměrná vzdálenost - centrální, mediánová, warolová(graf)
- 158) Diskriminační analýza**
- řeší problematiku vícerozměrné klasifikace
  - předmětem je nalezení statisticky nejvhodnějšího způsobu rozlišení mezi 2 či více soubory statistických jednotek
  - klasická úloha diskriminační analýzy spočívá v tom, že jsou předem známy 2 či více skupin jednotek a o každé víme do které skupiny patří
  - na základě naměřených údajů se pro každou skupinu vypočítá diskriminační fce sloužící k dodatečnému zařazování nových jednotek
  - do určité skupiny je zařazena ta jednotka, pro níž je pravděp. příslušnost této jednotky ke skupině největší
  - slouží k dodatečnému zařazení znaků do předem stanovených skupin
- 159) Faktorová analýza**
- umožňuje objasnit strukturu pozorovaných závislostí, redukuje počet výchozích proměnných pomocí hypotetických faktorů při minimální ztrátě informací a odhaduje skryté vztahy mezi proměnnými
  - použití při zpracování dotazníků
- 160) Kovarianční analýza**
- umožňuje rozčlenění variability proměnné  $y$  na části, které jsou přiřaditelné jednak kvalitativním a jednak kvantitativním vlivům
- 161) Analýza hlavních komponent**
- podstatou je transformace souboru napozorovaných proměnných do nových proměnných, tzv. hlavních komponent, které jsou vzájemně nezávislé a jsou seřazeny dle velikosti svého příspěvku ke vysvětlení celkového rozptylu napozorovaných proměnných
  - tato metoda citlivá na delta jednotek větší, než hodnoty normují
  - použití: odhalení skrytých vztahů mezi proměnnými
  - založena na bezzbytkovém vysvětlení celkového rozptylu proměnných

- 162) Kanonická korelační analýza**  
- vychází z logického členění proměnných do 2 skupin (výchozí + cílové veličiny) - každá skupina je charakterizována 1 soubornou veličinou = kanonickou proměnnou = lineární kombinace původních proměnných dané skupiny
- 163) Jak se provádí odhad pro celou regresní přímku?**
- 164) Rozdíl mezi jednoduchým a vícenásobným modelem regrese a korelace**  
- počet nezávisle proměnných  
- v modelu jednoduché stat. závislosti je předpokládáno, že změny závisle proměnných jsou vyvolány změnami jediné nezávisle proměnné, ostatní vlivy jsou považovány za náhodné  
- teorie vícenásobné regrese a korelace je zobecněním teorie jednoduché regrese a korelace
- 165) Korelace x regrese (2)**  
- korelace – těsnost, síla, míra závislosti mezi kvantitativními statistickými znaky  
- regrese – popis průběhu závislosti mezi kvantitativními znaky pomocí regresního modelu (tímto modelem je regresní funkce)
- 166) Základní podmínka metody nejmenších čtverců**  
-  $\sum d_i^2 = 0$  - vede k jednomu závěru  
-  $\sum d_i = \min$  - můžeme spočítat parametry
- 167) Metoda nejmenších čtverců a k čemu se používá**  
- matematická metoda s jejíž pomocí můžeme vypočítat parametry regresních fčí (s výjimkou fčí, které nemůžeme převést na aditivní tvar)
- 168) Pás spolehlivosti regresní přímky – vysvětlit (3)**
- 169) Jaká je regresní funkce nelineární v parametrech, jak parametry určujeme (3)**
- 170) Co se dá zjistit z regresní přímky?**
- 171) Co vyjadřuje regresní index a koeficient?**
- 172) Test regresního koeficientu.**
- 173) Podle jakých kritérií se vybírá nejvhodnější funkce pro vícenásobnou nelineární regresi?**
- 174) Postup při metodě stupňovité regrese (4)**
- 175) Napište rovnici regresní funkce a korelační koeficient, když máš zadáno x s pruhem (celkový průměr), y s pruhem (celkový průměr), s x, s y a s xy**
- 176) Lze testovat významnost regresní funkce a určit její interval spolehlivosti? Pokud ano, napište stručný postup.**
- 177) Určení závislosti u nelineární regrese**
- 178) Vícenásobná regresní funkce a jednoduchá regresní funkce – rozdíl**
- 179) Regrese na hlavních komponentách**  
- vícenásobný regresní model předpokládá nezávislost nezávisle proměnné  
- tento předpoklad bývá v praxi často nahrazen a proto se doporučuje do vícenásobné regrese nahradit nezávisle proměnnou hlavními komponentami
- 180) Regresní koeficient u lineární regrese**  
- byx - vyjadřuje o kolik se změní závisle proměnná y, jestliže se nezávisle proměnná x změní o jednotku  
- použití k odhadům změny
- 181) Aditivita vícenásobného regresního modelu**  
- funkci pro vícenásobnou regresi získáme jako součet jednoduchých regres. fčí
- 182) Jakým způsobem určujeme parametry regresních funkcí? (3)**  
- metoda nejmenších čtverců – obecná soustava normálních rovnic
- 183) Beta koeficienty a k čemu se používají**  
- normované přepočtení regresní koeficienty  
- k určení podílu jedn. nezávisle proměnné u regres. odhadu závisle proměnné. (vícenásobná lineární regrese)
- 184) Jak postupujeme při určení nejvhodnějšího typu regresní funkce**  
- na základě zkušenosti, logické posouzení, empirické posouzení (korel. pole), výpočet mnoha funkcí-výběr nejlepší

- 185) Jak provádíme výběr nejvhodnější nelineární regresní funkce? (3)**  
- podle toho, která nelineární regresní funkce má nejvyšší hodnotu indexu korelace
- 186) Co jsou to sdružené regresní přímky**  
- hodnoty korelačních koef. převádí v hodnoty z a vždy se řídí alespoň přibližně normálním rozdělením
- 187) Typy nelineárních regresních fci**  
- polynom, parabola 2. st., exponenciála, mocninná funkce, hyperbola, růstová funkce
- 188) Interakce ve vícenásobné regresní analýze (2)**  
- podíl nezávisle proměnných na regresním odhadu závisle proměnné  
- posuzuje vhodnost regresní funkce  
- představuje vzájemné působení proměnných (nejde to slovo přečíst)  
- vyjmenuj vztahy mezi koeficienty  
- ani jedna možnost
- 189) Korelační pole**  
- empirická metoda pro nalezení vhodné regresní funkce, vyjádření párové závislosti  
- vyjadřuje vzdálenost mezi teoretickou a skutečnou hodnotou u vybrané funkce
- 190) Korelační tabulka**  
- kombinační tabulka, která vyjadřuje průběh závislosti 2 proměnných (kvalitativní znaky)  
- na úhlopříčce je pevná závislost  
- čím jsou více soustředěné kolem úhlopříčky, tím je silnější závislost a naopak  
- těsnost závislosti v ní měříme korelačním poměrem
- 191) Korelační tabulka**  
- kombinační tabulka, která vyjadřuje průběh závislosti 2 proměnných (kvalitativní znaky)  
- těsnost závislosti v ní měříme korelačním poměrem
- 192) K čemu slouží Spearmanův koeficient korelace pořadových čísel**  
- charakterizuje těsnost jednoduché závislosti kvantitativních znaků  
- nabývá hodnot 0-1 původní hodnoty nahradíme pořadovými čísly je to u korel.  
- jedná se o neparametrickou charakteristiku  
- měří těsnost jakékoli statistické závislosti, která je monotónní  
- poskytuje rychlou a dostatečně přesnou informaci o těsnosti sledované závislosti
- 193) Totální koeficient korelace a determinace**  
- totální koeficient korelace - vyjadřuje těsnost závislosti Y na všech nezávislých prom. X  
- totální koeficient determinace - druhá mocnina korelace; udává z kolika % je nez. proměnná Y ovlivněna všemi uvažovanými nezávisle proměnnými X
- 194) Koeficient vícenásobné totální korelace**  
- měří těsnost závislosti závisle proměnných na všech uvažovaných nezávisle proměnných
- 195) Koeficient parciální korelace**  
- měří těsnost závisle proměnných na jedné z uvažovaných nezávisle proměnných
- 196) Korelační index je (možnosti) (2)**  
- měříme jím těsnost závislosti u nelineárních funkcí  
- je v intervalu od 0 do 1, čím více se blíží 1 tím je závislost silnější, čím blíže 0 tím je závislost slabší
- 197) Jak ověřujeme statistickou významnost korelačního indexu a korelačního koeficientu? (3)**  
- korelační koeficient - podle t-Studentova rozdělení, tabulky kritických hodnot korelačního koeficientu, neprovádíme test, pouze porovnáváme hodnotu korel. koeficientu s tabulkovou, je-li vypočtená > tabulková, můžeme rovnou korel. koeficient považovat za statisticky významný  
- korelační index - netestujeme
- 198) Co měří dílčí korelační koeficient a co vícenásobný korelační koeficient?**  
- dílčí korelační koeficient vyjadřuje těsnost lineárního vztahu mezi dvěma proměnnými při vyloučení vlivu jedné nebo více dalších proměnných  
- vícenásobný koeficient korelace charakterizuje těsnost závislosti jedné proměnné na lineární kombinaci jiných proměnných
- 199) Koeficient determinace -  $r^2$ ; vyjádření v %**  
- udává z kolika % je závisle proměnná ovlivněna uvažovanou nezáv. Proměnnou (u lineární regrese a korelace)
- 200) Totální koeficient determinace**  
- vyjadřujeme v % a udává nám z kolika % je závislá proměnná ovlivněna uvažovanými nezávisle proměnnými
- 201) K čemu lze prakticky využít determinační index? O čem na jeho základě můžeme rozhodnout?**  
- index determinace vynásobený 100 udává relativně v procentech tu část rozptylu závisle proměnné z, kterou se podařilo vysvětlit použitou regresní funkcí  
- čím bližší je jedné, tím je daná závislost silnější.
- 202) Může být vícenásobný korelační koeficient menší než některý z jednoduchých korelačních koeficientů?**  
- nemůže <= vícenásobný je složen z jednoduchých

- 203) **Jakou metodu byste použili k posouzení, jestli je prodej určitého zboží v průběhu dne rovnoměrně rozložen?**  
 - rozptyl nebo řetězový index
- 204) **Posuďte, která varianta vztahů mezi koeficienty je možná:**  
 a)  $b_{yx}=-5$  ,  $b_{xy}=-0,2$   
 b)  $r_{yx}=-0,72$  ,  $b_{xy}=-2,1$   
 c)  $b_{yx}=0,5$  ,  $b_{xy}=2,1$   
 d)  $b_{yx}=-0,1$  ,  $b_{xy}=0,7$   
 Možná je varianta A - Nesmí mít opačná znaménka a jejich součin musí být menší nebo roven jedné.
- 205) **Je možné, aby se korelační koeficient rovnal korelačnímu poměru? Kdy?**  
 - ano. Když je daná závislost naprosto shodná se svou regresní přímkou.
- 206) **Na jakém principu je založena konstrukce měř těsnosti závislosti?**  
 - na principu rozkladu rozptylů
- 207) **U kterých z uvedených regresních funkcí lze k odhadu parametrů použít metodu nejmenších čtverců: přímka, parabola, exponenciála ? Názor vysvětlete.**  
 - u přímky a paraboly, protože jsou lineární v parametrech na rozdíl od exponenciály, která není - tam je nutná lineární transformace.
- 208) **Jak určíme ve vícenásobné lineární regresní funkci podíl jednotlivých nezávisle proměnných v regresním odhadu závisle proměnné**  
 -  $\beta$  koeficienty regresní fce (dílní regresní koeficienty na které jsou přepočítány nezávislé odchylky)- o kolik směrodatných odchylek se změní závisle proměnná jestliže se nezávisle proměnná změní o 1 odchylku
- 209) **Co je multikolinearita a jak ji zjistíš? (5)**  
 - je to závislost mezi vysvětlujícími proměnnými  
 - informace o ní čerpáme z matice korelačních koeficientů a jejího determinantu  
 - jestliže jsou párově nekorelované, multikolinearita neexistuje a  $D=1$ , pokud  $D=0$ , potom mluvíme o úplné multikolinearita  
 - většinou je škodlivá od hodnoty 0,75  
 - jestliže multikolinearita existuje, tak je některá proměnná x v modelu zbytečná
- 210) **Co je zdánlivá korelace?**  
 - zdánlivá korelace spočívá v tom, že je někdy možné pozorovat silnou závislost mezi proměnnými i v případě, kdy mezi proměnnými ve skutečnosti závislost buď skoro, nebo vůbec neexistuje.
- 211) **Doprovodná regrese**  
 - u analýzy určujeme, zda kromě regrese na kvantitativním znaku existuje i regrese na kvalitativním znaku - f-test
- 212) **Čím měříme těsnost závislosti**  
 - u lineárních závislostí - koef korelace, determinace  
 - u nelineárních závislostí - index korelace, determinace
- 213) **Kanonická korelační analýza**  
 - vycházejí z logického členění proměnných do dvou skupin  
 - každá skupina proměnných je charakterizována 1 souhrnnou veličinou, tzv. kanonickou proměnnou = lineární kombinace pův. prom. dané skup.
- 214) **Jak zjistíme zda je rozdíl korelačních koeficientů statisticky významný**  
 - tabulky - pomocí testu o korelačním koeficientu  
 -  $tr = r_{yx} / (\text{odm}(1-r^2 \text{ na } 2xy) / (n-2))$   
 - tr - má studentovo rozdělení pro n-2 stupňů volnosti  
 - t větší než t alfa - statisticky významný
- 215) **z-transformace, kdy ji používáme, její význam**  
 - hodnoty  $r_1$  převádíme na  $z_1$  a  $r_2$  na  $z_2$  u testu o rozdílu 2 korelačních koeficientů  
 - metoda pro intervalový odhad korelačního koeficientu, jestliže  $n < 100$   
 - při intervalovém odhadu korelačního koef. v případě že n je menší než 100 není možná normální aproximace
- 216) **Fischerova Z-transformace**  
 - speciální postup, kdy hodnoty korelačního koeficientu převádíme (pomocí tabulky) na hodnotu  $Z(r_1=Z_1, r_2=Z_2)$   
 - testovací kritérium -  $u = (Z_1/Z_2) / (\text{odm}(1/(n_1-3)) + (1/(n_2-3)))$   
 - u je větší než u alfa ...  $H_0$  se zamítá
- 217) **Koeficient korelace – co vyjadřuje, jaké známe (4)**
- 218) **Jak určíme parametry mocninné funkce**
- 219) **Kdy se používá korelační poměr, jakých nabývá hodnot**
- 220) **Rozdíl mezi dílčím a totálním korelačním koeficientem**

- 221) **Rozdíl mezi volnou a pevnou závislostí**
- 222) **Index korelace – co to je, co vyjadřuje, jak je vymezen, kdy se používá**
- 223) **Jakých hodnot nabývá korel. index a jak ho interpretujeme; co vyjadřuje index determinace**  
 - korelační index nabývá hodnot  $<0;1>$ , čím více se blíží 1 tím je závislost silnější, čím blíže 0 tím je závislost slabší  
 - index determinace - I na 2xy - vyjádřen v % udává z kolika % je závisle proměnná ovlivněna uvažovanou nezávisle proměnnou
- 224) **Uveďte vztah mezi korelačním koeficientem a regresními koeficienty –  $ryx = \sqrt{byx.bxy}$**
- 225) **Čím měříme těsnost závislosti**  
 - korelačním koeficientem (lineární regrese), korelačním indexem (nelineární regrese)
- 226) **Jak se vyjadřuje doprovodná regrese u kovariance?**
- 227) **Je možné, aby se korelační koeficient rovnal korelačnímu poměru? Kdy?**  
 Ano. Když je daná závislost naprosto shodná se svou regresní přímkou.
- 228) **Množné znaky - kontingenční tabulka**  
 - alespoň 1 znak většinou více než 2, měříme pouze test závislosti; Pearsenův koeficient kontingence vyzívá chí-test
- 229) **K čemu slouží normovaný koeficient kontingence**  
 - kontingenční tabulky mají různý počet řádků a sloupců - používáme ho pro porovnávání
- 230) **Koeficient asociace slouží k:**  
 - vyjádření těsnosti 2 alternativních znaků
- 231) **Čím určujeme těsnost závislosti v asociační tabulce**  
 - koeficient asociace  $<-1;1>$ , obdoba koeficientu korelace
- 232) **Asociační tabulka**  
 - vyjadřuje vztah 2 alternativních znaků  
 - asociační přímkou, těsnost, koeficient asociace (Yuelův koeficient asociace) v tabulce prvky a a b b  
 - používají se zde upravené chí-testy
- 233) **Jaké funkce popisují průběh závislosti v asociační tabulce**  
 - lineární fce
- 234) **Jak měříme těsnost závislosti v kontingenční tabulce**  
 - Pearssonův koef kontingence - C (0,1)  
 - čupronův koef kont. - K (0,1)  
 - normovaný koef kontingence - Cn (0,1)  
 - vyjádřit průběh závislosti v kontingenční tabulce neumíme
- 235) **Jak ověřujeme závislost znaků v asociační tabulce**  
 - chí-kvadrát test, fischerův test
- 236) **Jaké funkce lze použít k popisu průběhu závislosti v asociační tabulce**  
 - pouze lineární asoc. přímkou
- 237) **Rozdíl mezi korelační a kontingenční tabulkou?**
- 238) **Kontingenční tabulka, nakreslit obecné schéma**
- 239) **Fischer-pravděpodobnostní test tabulek**
- 240) **K čemu slouží normovaný koeficient kontingence**  
 - kontingenční tabulky mají různý počet řádků a sloupců - používáme ho pro porovnávání  
 - aby byl dostačující - postačující odhad
- 241) **Jaké testy se používají v asociační tabulce**  
 - upravený chí-kvadrát test; fisherův test  
 - jakými způsoby lze vyjádřit trend časové řady  
 - graficky, mechanicky (klouzavé prům.), analyticky (trend. fce.)
- 242) **Charakteristické rysy časových řad**  
 - uspořádaná řada údajů, které odlišují v čase  
 - rysy – trend (vývojové tendence v čr-vzestupný, sestupný, stacionární), kolísání (odchylky od rovnoměrného vývoje-periodické(cyklické, sezónní), náhodné(okamžikové, intervalové))
- 243) **Elementární charakteristiky čas. Řad**  
 - slouží k rychlé informaci o charakteru a chování ukazatele v čas. řadě
- 244) **Harmonická analýza, podstata, k čemu se používá a kde**  
 - soubor postupů používaných při analýze vlnitých pohybů u přírodních jevů

- v některých ČŘ se přesně pravidelně odchylojí skutečné hodnoty od celkové tendence
  - graf připomíná spojitě periodické funkce goniometrického typu
- 245) Co je autokorelace? Kde se s ní setkáváme a jak ji měříme?**
- je korelace mezi sousedními odchylkami od trendu
  - ověřujeme ji pomocí Durbin- Watsonova testu
  - závislost mezi 2 po sobě jdoucími členy v čř
  - u takové čř nelze provést měření těsnosti závislosti
  - zjišťuje se pomocí koeficient autokorelace (nízké hodnoty - neexistuje autokorelace)
- 246) Jak se měří průměr v okamžikové a intervalové ČŘ?**
- 247) Analýza časových řad**
- trend x periodická složka x náhodná složka
- 248) Kdy používáme klouzavé průměry**
- při mechanickém popisu trendu - očišťuje čř od náhodného a period. kolísání
- 249) K čemu používáme klouzavé průměry?**
- k vyrovnávání ČŘ
  - nevyrovnáme ji celou najednou, ale po částech
- 250) Co si představujete pod pojmem "dekompozice časové řady"?**
- rozklad čas. řady na 4 složky: trend, sezónní kolísání, cyklické kolísání a náhodná složka
  - u aditivního modelu to sčítáme, u multiplikativního násobíme
- 251) Jaké jsou předpoklady o náhodné složce časové řady? Proč ověřujeme jejich splnění a co to znamená, když se přesvědčíme, že nejsou splněny?**
- kompenzace v rámci časové řady, homoskedasticita (jejich variabilita se v čase nemění), jsou vzájemně nekorelované
  - nesplnění = nepodařilo se v časové řadě eliminovat systematickou složku beze zbytku
- 252) Co je extrapolace časové řady a jaký má význam?**
- děláni prognóz do budoucnosti
  - nejčastější způsob použití extrapoláčnických kritérií je založen na simulaci spočívá v tom, že z analýzy řady oddělíme určitou část pozorování a na vhodný T funkce usuzujeme podle toho, jak dobře extrapoluje tato pozorování.
- 253) Pomocí jakých měř posoudíte přesnost extrapolace?**
- Index korelace a determinace, u lin. záv. pomocí korel. koeficientu. Nebo pokud nemůžu model popsat žádnou závislostí tak poměrem determinace.
- 254) Jak vyberete vhodnou metodu pro extrapolaci časové řady?**
- nakreslím graf a posoudím průběh, potom analyzuji růstové charakteristiky a difference
- 255) Z údajů časových řad dvou ukazatelů y a x byl vypočten korelační koeficient  $r_{yx} = 0,8$ . Jaké závěry z této hodnoty můžeme učinit?**
- mezi y a x existuje středně silná lineární závislost
- 256) Z údajů časových řad ukazatele y a ukazatele x byl vypočten korelační koeficient  $r_{yx} = 0,6$ . O čem vypovídá?**
- Středně silná závislost
- 257) Časová řada čtvrtletních údajů je rostoucí a sezónní indexy rostou úměrně trendu. Jak byste posoudili sezónnost takové časové řady?**
- jedná se o proporcionální sezónnost
- 258) Co chápeme pod pojmem paralelismus časových řad a co způsobuje?**
- Souběh časových řad: Systematické složky časových řad, zejm. celkový vývoj tendence, mohou mít velmi podobný průběh v čase, což může vést k pozorování silné závislosti mezi proměnnými, jež ve skutečnosti vůbec závislé nejsou. Proto je třeba zjistit, zda existuje nějaký vztah mezi náhodnými složkami, a pak teprve pak usuzovat na souv, mezi ČŘ.
- 259) Popište základní princip exponenciálního vyrovnávání časové řady.**
- Pozorování blízká současnosti jsou pro odhad parametrů důležitější, než pozorování vzdálenější, starší, a měla by jim proto být přisuzována větší váha.
- 260) Jaké znáte metody pro získání odhadů parametrů trendových funkcí nelineárních v parametrech?**
- linearizovat model vhodnou transformací, pak požadavek kompenzace kladné a záporné odchylky empirických hodnot od hodnot vyrovnaných a metoda nejmenších čtverců (aby součet čtverců popsanych odchylek byl minimální)
- 261) Náhodné kolísání časové řady, jak ho měříme**
- výsledek náhodných jevů, které nelze předurčit
  - pomocí abs. prům. odch. a rel. prům. odch.
- 262) Jak se určí průměr časové řady okamžikové a intervalové**
- okamžiková:  $y = \frac{y_1 + y_2 + \dots + y_{n-1} + y_n}{n}$
  - intervalová - počítá se jako prostý ar. průměr
- 263) Jakými zp. lze vyjádřit trend časové řady**
- graficky, mechanicky (klouz. prům.), analyticky (trend. fce.)
- 264) Jak počítáme v čas. řadě koef. růstu a prům. koef. růstu**
- koef. růstu -  $K_i = \frac{y_i}{y_{i-1}} \cdot 100$
  - prům. koef. růstu -  $k = \sqrt[n]{k_1 \cdot k_2 \cdot \dots \cdot k_n}$

- 265) Sezónní index je**  
- poměr skutečných a teoretických hodnot
- 266) Extrapolace**  
- jedná se o statistické prognózování kdy pomocí trendové fce a sezónních indexů je možno odhadnout budoucí vývoj daného ukazatele
- 267) Interpolace**  
- přibližné určení chybějící hodnoty sledovaného ukazatele v ČR za předpokladu, že známe sousední hodnoty  
- provádí se 2 způsoby: prostřednictvím 2 sousedních údajů; prostřednictvím všech hodnot v ČR
- 268) Popište základní princip exponenciálního vyrovnávání časové řady.**  
- Pozorování blízká současnosti jsou pro odhad parametrů důležitější, než pozorování vzdálenější, starší, a měla by jim proto být přisuzována větší váha.
- 269) Co jsou odvozené časové řady a aspoň dva typy**  
(z druhotných ukazatelů, např. HDP/na obyvatele)
- 270) Čím vyjadřujeme vliv sezónního kolísání**
- 271) Co je stacionární časová řada**
- 272) Odvozené časové řady**
- 273) Postup při korelaci časové řady.**
- 274) Co jsou bazické a co řetězové indexy a jaký je mezi nimi vztah?**  
- jsou to indexy, které se týkají časové posloupnosti  
- bazický index má základ v zákl. období, řetězový vždy v předchozím období  
- bazické se dají konstruovat pomocí násobení řetězových a řetězové zase dělením bazických
- 275) Jaký index (indexy) použijete k posouzení změny cen vyráběné produkce u podniku, který vyrábí 7 různých výrobků?**  
- Souhrnný index - různé druhy výrobků nemůžeme sčítat - Laspeyresův, Paascheho, Fischerův...
- 276) Která znáte rozdělení spojitých náhodných veličin? Uveďte oblasti jejich použití.**  
Normální, Logaritmicko-normální, Exponenciální, Chi-kvadrát, Studentovo, Rozdělení F
- 277) Jaká je hlavní myšlenka rozkladu indexů metodou postupných změn?**  
Index hodnot analyzovaného ukazatele rozkládáme na dílčí analytické indexy, které se skládají součinem.
- 278) Může nastat situace, že index ceny hovězího masa stoupne, i když ve všech prodejnách cena poklesla? Zdůvodněte svůj názor.**  
Z ekonomického hlediska ano. Asice v případě, že by indexy cen ostatního zboží a služeb stouply více.
- 279) Jak budete postupovat, chcete-li zjistit, jak se na změně průměrné dovozní ceny určitého výrobku podílely změny cen z jednotlivých dovozních zemí a jak změna struktury dovozu podle dovozních zemí?**  
Indexem proměnlivého složení
- 280) Charakterizujte situaci, kdy použijete složený individuální cenový index a kdy souhrnný cenový index.**  
Složený index: Chceme-li zjistit, jak se na změně průměrné ceny určitého výrobku podílely změny jednotlivých cen, a jak změna struktury  
Souhrnný index: nemůžeme-li sčítat různé druhy výrobků - Laspeyresův, Paascheho, Fischerův...
- 281) Podle jakých hledisek dělíme indexy**  
- věcné, časové, prostorové
- 282) Sezónní index**  
- podíl pův. hodnot a hodnot očištěných od sez. vlivů
- 283) Souhrnné indexy úrovně - (2)**  
- Indexy stejnorodý intenzity ukazatele; nestejnorodý extenzity ukazatele; srovnávající nest. prod.; nest. intenzity ukazatele; ani jedna
- 284) Výsledný ukazatel intenzity je stejnorodý pokud:**  
- oba intenzivní ukazatele jsou stejnorodé; oba extenzivní ukazatele jsou stejnorodé; oba nestejnorodé; jeden extenziv. uk. je stejnorodý, druhý nestejnorodý
- 285) Stejnorodé a nestejnorodé ukazatele**  
- stejnorodé - lze-li srovnat ext. ukaz. součtem v přír. měrných jedn.  
- nestejnorodé - nelze je mezi sebou spočítat
- 286) Indexy**  
- podíl dvou hodnot téhož ukazatele; individuální a souhrnný; věcný, prostorový, časový  
- individuální - indexy stejnorodých ukazatelů  
- souhrnné - indexy nestejnorodých ukazatelů  
- obě tyto skupiny se dělí na indexy množství a indexy úrovně

- 287) **Typy stat. ukazatelů**  
- okamžikové, intervalové, primární, sekundární, extenzivní, intenzivní, stejnorodé, nestejnorodé
- 288) **Rozklad indexu hodnoty.**
- 289) **Rozdíl mezi souhrnným indexem množství a indexem úrovně.**
- 290) **Fischerův index – k čemu se používá.**
- 291) **Montgomeryho index – k čemu (3)**
- 292) **Stejnorodé a nestejnorodé ukazatele**  
- stejnorodé - lze li srovnat ext. ukaz. součtem v přír. měrných jedn.  
- nestejnorodé - nelze je mezi sebou spočítat
- 293) **K čemu Guttmanův koeficient prediktability (3)**
- 294) **Co je to vydatnost?**
- 295) **K čemu využíváme Theilův koeficient?**
- 296) **Co vyjadřuje Bayesův vzorec?**  
- v případě, že jsou známy nejen nepodmíněné pravděpodobnosti  $P(B_i)$ , ale i podmíněné pravděp.  $P(A/B_i)$  a je známo, že výsledkem je nastoupení jevu A, lze  $P(B_i)$  spočítat podle vzorce:  $P(B_i) = P(B_i) * P(A/B_i) / \sum P(B_i) * P(A/B_i)$
- 297) **Zemědělský závod**  
- souhrn pozemků větších než 0,1 ha na kterých se hospodaří  
- souhrn pozemků menší než 0,1 ha když:
- 298) **Jak odhadujeme v zemědělství ha výnosy**  
- subjektivní odhad - celé území se rozdělí na obvody a v každém se odhaduje prům. výnos  
- objektivní metoda - výběrové šetření, přímé měření; sledují se ha výnosy - klouz. prům.
- 299) **Kultura**  
- pozemek s trvale určeným způsobem využívání půd (orná půda, vinice), plodina - brambory rané...
- 300) **Rozdíl mezi sklizňovou plochou a sklizenou plochou.**
- 301) **Statistická síť pro zemědělství.**
- 302) **Co je šetření Agrocenzus a kdy se toto šetření naposled uskutečnilo**
- 303) **Systém rodinných účtů**
- 304) **Vyjmenovat sčítací testy**
- 305) **Agrocenz**  
- vychází z prahových hodnot; výběrová šetření; vysoké náklady
- 306) **Jaká měření provádí statistika v zemědělské rostlinné výrobě**