

Subiecte finale

Verificare an II Biochimie 1998-99

Prezentăm în continuare un exemplu de subiecte din verificarea finală aplicată anului II Biochimie 1998-99, la prima examinare. Notele obținute, împreună cu notele primite la fiecare lucrare practică, au produs distribuția analizată în problemele 1 și 6 din Lp 9. (Vezi și problema 1 din Lp 4.)

Proba I (TG - Durata 5')

1. Variabilele cantitative reprezintă:

1. variabile ale căror valori pot fi ordonate sau măsurate.
2. variabile ale căror variante pot fi ordonate sau măsurate.
3. variabile ale căror variante pot fi clasate, ordonate sau măsurate.

2. Alegeți definiția corectă pentru mediana unui șir:

1. O valoare care împarte șirul ordonat în două subșiruri de volume egale.
2. Valoarea care împarte seria statistică în două subserii de volume egale.
3. O valoare care împarte șirul ordonat în două subșiruri egale.

3. Ce afirmație este corectă:

1. Nivelul de semnificație = puterea testului.
2. Nivelul de semnificație = riscul de speța II.
3. Nivelul de semnificație = gradul de încredere
4. Nivelul de semnificație = 1 - gradul de încredere

4. Alegeți afirmația corectă:

1. Precizia unei estimații se poate îmbunătăți experimental prin randomizare și volum cât mai mare.
2. Precizia unei estimații se poate îmbunătăți experimental printr-un volum cât mai mare.
3. Precizia unei estimații se poate îmbunătăți experimental prin randomizare.

5. Dacă un pacient obține rezultate pozitive la două teste cu specificitate Sp_1 , respectiv Sp_2 , este ca și cum i s-ar fi aplicat un singur test care are:

1. $Sp = Sp_1 + Sp_2 - Sp_1 \cdot Sp_2$.
2. $Sp = Sp_1 \cdot Sp_2$.
3. $Sp = Sp_1 + Sp_2$.

6. Pentru o tabelă de contingență cu număr de linii egal cu numărul de coloane ($p = q$), coeficientul de contingență al lui Ciuprov, T , are valorile cuprinse între:

1. 0 și 1.
2. -1 și 1.
3. $-\infty$ și $+\infty$.

7. Alegeți afirmația ERONATĂ:

Coeficientul de corelație a rangurilor al lui Spearman, R_S , are valoarea 1 dacă și numai dacă:

1. corelația este absolută.
2. rangurile sunt identice.
3. corelația este perfectă.

Proba a II-a (TC – Durata 10')

1. Cea mai importantă proprietate a mediei este _____.
2. Când obținem $\alpha_c < 0,001$ înseamnă că testul a evidențiat o diferență _____.
3. Dintre indicatorii de împrăștiere, _____ este un indicator bazat pe indicatori de tendință extremă, iar _____ este bazat pe indicatori de tendință intermediară.
4. _____ testului se notează cu π și este egală cu _____, în care _____ este eroarea de speța a II-a.
5. Compararea unei medii cu o valoare teoretică este un test de _____.
6. Care este răspunsul corect și complet la întrebarea: intervalul (20; 23) de încredere 90% acoperă sau nu media populației ?
 - a) intervalul (20; 23) acoperă media în 90% din cazuri.
 - b) intervalul (20; 23) acoperă media în 90% din cazuri, iar în 10% din cazuri nu o acoperă.
 - c) nu știm.

182 Dragomirescu L. Lucrari practice de biostatistica. Editia a III-a revazuta si adaugita, 263 pp. Editura "Agronomica", Bucuresti, 2003.

d) nu știm, dar presupunem că acest interval face parte din cele 90% intervale de forma $(m - 1,28 \cdot S(m); m + 1,28 \cdot S(m))$ care acoperă media populației, mai degrabă decât din cele 10% care nu o acoperă.

7. Când σ este cunoscut și $n > 30$, distribuția mediei de eșantionaj poate fi considerată _____.

8. Dacă H_0 este adevărată, dar o respingem, testul produce o eroare _____.

9. Dacă 0,1515-cuantila unilaterală inferioară este -1,03, atunci 0,3030-cuantila bilaterală inferioară este _____.

10. Se știe că sub distribuția normală standard aria la stânga lui 3 este 0,9987. Cât este aria la dreapta lui $\mu + 3\sigma$ sub o distribuție normală de medie μ și abatere standard σ ?

R: _____.

Proba a III-a (Exerciții sau probleme – Durata 30')

(Se recomandă utilizarea unui calculator de buzunar cu funcția radical și se permite consultarea volumului "Biostatistică pentru începători")

Observație: Studenții au fost avertizați că primesc trei probleme din care una cere explicit doar o rezolvare de principiu, iar una din celelalte două este nerezolvabilă, fără a se preciza însă care este aceasta.

1.

Estimați media lungimii corpului unui individ într-o populație de *Cyprinus carpio*, printr-un interval de încredere 90%, dacă 9 indivizi, extrași aleator din aceasta, au următoarele lungimi măsurate în centimetri: 25 28 20 35 25 27 30 28 27. Rotunjiți rezultatele la două zecimale.

2.

Pentru a determina efectul hipotensiv al unui nou medicament în tratamentul afecțiunilor cardiovasculare în stadii incipiente, s-a luat în studiu un lot de 236 pacienți. Într-o primă etapă, li s-a administrat subiecților placebo. După o lună,

li s-a administrat o doză unică din noul medicament de testat. S-au obținut următoarele rezultate:

<i>Preparatul testat</i>	<i>Efect hipotensiv</i>	<i>Nici un efect</i>
Placebo	189	47
Noul medicament	198	38

Stabiliți dacă noul medicament este eficace sau nu.

3.

În cadrul unui studiu privind impactul poluării cu cupru asupra lucrătorilor din industria cuprului, s-au prelevat probe de urină de la un eșantion aleator de 35 de lucrători. Apoi s-a dozat cuprul în fiecare probă, în vederea comparării cu concentrația maximă admisă în literatură pentru acest element. Stabiliți tipul (din toate cele trei puncte de vedere) de test ce trebuie aplicat și formulați ipoteza nulă și cea alternativă.

Teste de biostatistică - selecții finale loturi olimpice

Următoarele două grupaje de câte 10 întrebări au făcut parte din selecțiile finale ale loturilor olimpice de biologie în anii 1998 și 1999. Din motive lesne de înțeles le prezentăm fără rezolvări.

Anul 1998. Durata 10'

1. Cinci bacterii dintr-o cultură au diametrele:

5 3 6,87 2

Mediana acestor diametre este:

- a. 6,8.
- b. 5.
- c. 5,9.
- d. 4,76.

2. Media diametrelor de la punctul anterior este:

- a. 5.
- b. 5,9.
- c. 6,5.
- d. 4,76.
- e. 6,8.

3. Dispersia necorectată a diametrelor de la punctul 1, rotunjită la o zecimală, este:

- a. 4.
- b. 3,98.
- c. 5.

d. 4,988.

4. Abaterea standard (necorectată) s , a diametrelor de la punctul 1, este aproximativ:

- a. 1,997.
- b. 2,236.
- c. 2.
- d. 2,22356.
- e. 3.

5. Eroarea standard a mediei (de eşantionaj) $S(m)$, în cazul seriei de la punctul 1, este:

- a. 1 deoarece $S(m) = s / \sqrt{n-1}$.
- b. 1 deoarece $S(m) = s / \sqrt{n}$.
- c. 1,118 deoarece $S(m) = s / \sqrt{n-1}$.
- d. 1,12.

6. Considerând că bacteriile de la punctul 1 au fost extrase aleator dintr-o cultură, un interval de încredere 95% pentru media diametrelor în cadrul întregii culturi va fi:

- a. $4,76 \pm 2,776$.
- b. $5 \pm 2,776 \cdot 1,12$.
- c. $5,9 \pm 2,776 \cdot 1,118$.
- d. $5 \pm 2,776 \cdot 1,118$.

unde cuantila 2,776 s-a extras din tabela Student pentru $n-1$ grade de libertate ($n =$ volumul eşantionului) și coeficientul de risc 5% plasat bilateral.

7. Pentru a se verifica dacă două recipiente conțin aceeași cultură de bacterii s-a extras din fiecare câte o probă și s-au măsurat diametrele celor 30, respectiv 50 bacterii din cele două probe. Știind că în cazul distribuției normale standard valoarea 1,96 lasă la dreapta sa 0,025 din aria de sub curbă și că s-au obținut mediile 4μ , respectiv 8μ , iar dispersiile corectate sunt $60\mu^2$, respectiv $100\mu^2$, ce test trebuie aplicat pentru a se face verificarea din enunț:

- a. Testul t de conformitate.
- b. Testul t pentru observații perechi.
- c. Testul Student (t) pentru observații independente.
- d. Testul t de comparare a unei medii empirice cu o medie teoretică.

8. Care este statistica testului care trebuie aplicat în problema anterioară:

$$\text{a. } t = \frac{m - \mu}{s / \sqrt{n-1}} \quad \text{b. } t = \frac{m_1 - m_2}{\sqrt{\frac{s_{c_1}^2}{n_1} + \frac{s_{c_2}^2}{n_2}}} \quad \text{c. } t = \frac{m_1 - m_2}{\frac{s_{c_1}^2}{n_1} + \frac{s_{c_2}^2}{n_2}}$$

9. Ce decizie statistică se obține:

- Acceptăm ipoteza nulă cu riscul 5%.
- Respingem ipoteza nulă cu riscul 5%.
- Acceptăm ipoteza nulă cu riscul de a o respinge de 5%.
- Respingem ipoteza nulă.

10. Cum interpretăm rezultatul statistic obținut la punctul anterior:

- Considerăm, cu riscul 5%, că recipientele conțin culturi de bacterii diferite.
- Considerăm, cu riscul 5%, că recipientele conțin aceeași cultură.
- Considerăm că recipientele conțin aceeași cultură.
- Considerăm că recipientele conțin aceeași cultură sau că nu am măsurat suficiente bacterii pentru a detecta o eventuală deosebire a diametrelor.

Anul 1999. Durata 10'

(Se va utiliza un calculator de buzunar cu funcția radical. Se permite consultarea lucrării "Biostatistică pentru începători".)

De o parte și de alta a unui râu a fost plantată cânepă din același soi. La maturitate s-a extras câte un eșantion aleator din fiecare lot pentru a se testa dacă cele două loturi conțin plante cu înălțimi medii diferite sau nu. Din primul lot s-au extras 12 plante, iar din al doilea 10. Mediile lungimilor plantelor au fost 85 cm, respectiv 81 cm, iar dispersiile corectate 16 cm², respectiv 25 cm².

1. Alegeți ipoteza nulă adecvată acestei probleme:

- Diferența între cele două medii este semnificativă.
- Diferența între medii este accidentală, datorându-se doar fluctuațiilor de eșantionaj.
- Diferența între medii se datorează condițiilor diferite din cele două loturi.

2. Alegeți ipoteza alternativă adecvată, cel mai corect și complet formulată:

- Diferența între medii este semnificativă, dar între dispersii nu este semnificativă.

- b. Diferența între medii nu a apărut accidental, ci se datorează unei anumite cauze (de exemplu, condiții diferite oferite de cele două terenuri sau tratamente diferite ale celor două loturi, etc.)
- c. Diferența între medii nu a apărut accidental, ci se datorează condițiilor de mediu diferite oferite de cele două terenuri.
3. Ce test statistic trebuie aplicat pentru rezolvarea acestei probleme:
- Testul t bilateral pentru observații independente.
 - Testul χ^2 .
 - Testul t unilateral dreapta pentru observații independente.
 - Testul t bilateral pentru observații perechi.
4. Făcând notațiile următoare pentru cele două eșantioane:

Volume:	Medii:	Dispersii (corectate):
$n_1 = 12$	$m_1 = 85 \text{ cm}$	$s_{c1}^2 = 16 \text{ cm}^2$
$n_2 = 10$	$m_2 = 81 \text{ cm}$	$s_{c2}^2 = 25 \text{ cm}^2$

și aplicând formula dispersiei comune:

$$s_p^2 = \frac{(n_1 - 1)s_{c1}^2 + (n_2 - 1)s_{c2}^2}{n_1 + n_2 - 2},$$

se obține pentru s_p^2 valoarea:

- 22,78.
 - 30,14.
 - 20,05.
 - 1,23.
5. Pentru aplicarea testului necesar se calculează statistica

$$t = \frac{m_1 - m_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

care are valoarea rotunjită la o zecimală:

- 1,2.
- 1,71.
- 2,7.
- 2.

Decizia statistică o vom lua consultând tabela corespunzătoare testului stabilit care conține următoarele valori critice asociate nivelelor de semnificație uzuale:

Nivele de semnificație:	0,1	0,05	0,01	0,001
Valori critice (tabelate):	1,725	2,086	2,845	3,850

6. În concluzie putem respinge ipoteza nulă pentru (sau cu riscul p):
- $0,01 < p < 0,05$.
 - $0,001 < p < 0,01$.
 - $0,05 < p < 0,1$.
 - $p < 0,001$.
7. Luând drept valoare standard pentru nivelul de semnificație valoarea 0,05, să se specifice dacă ipoteza nulă a fost respinsă:
- da.
 - nu.
 - nu știm.
8. Cum se interpretează statistic decizia anterioară:
- Ipoteza nulă poate fi adevărată sau nu avem suficiente date pentru a fi, eventual, respinsă.
 - Ipoteza nulă este adevărată sau nu avem suficiente date pentru a fi respinsă.
 - Datele nu sunt compatibile cu ipoteza nulă, ci cu cea alternativă.
9. Cum se interpretează biologic decizia statistică efectuată:
- Condițiile de pe cele două terenuri au fost aceleași.
 - Este posibil ca plantele de pe cele două loturi să aibă înălțimile medii egale sau nu avem suficiente date pentru a susține, eventual, contrariul.
 - Plantele din cele două loturi au înălțimi medii diferite semnificativ.
10. În general, în cazul unui test statistic, dacă respingem ipoteza nulă:
- înseamnă că este adevărată ipoteza alternativă.
 - nu vom ști niciodată care dintre cele două ipoteze este adevărată.
 - nu știm care ipoteză este adevărată decât dacă cercetăm în întregime populația sau populațiile statistice (ceea ce în cele mai multe cazuri este sau foarte greu sau imposibil).
 - înseamnă că este improbabilă și, deci, este falsă.

Verificare anul I Biologie, septembrie 2002

[Cu carte și calculator, cca. 30 minute]

1. Pentru următoarea serie statistică:

4 4 3 7 4 5 4 6

- să se calculeze: moda, cuartila inferioară, mediana, cuartila superioară, media, amplitudinea, dispersia, abaterea standard, coeficientul de variație.

- b. să se calculeze cuartila inferioară de ordinul 10 și decila superioară de ordinul 100.
- c. să se determine gradul de concentrare/ dispersare al distribuției.
2. Diametrul trunchiului la arbori este o variabilă _____(a)_____ care după mulțimea de reprezentare este variabilă _____(b)_____.
3. Specia este o variabilă _____(a)_____ ce se poate reprezenta pe scala _____(b)_____.
4. Gândind cantitativ variabilitatea este concepută ca o _____(a)_____ iar gândind calitativ variabilitatea „devine” _____(b)_____.

5. Fie distribuția de abundențe absolute:

$$\begin{pmatrix} x_j \\ N_j \end{pmatrix} = \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \\ 2 & 3 & 4 & 6 & 7 \end{pmatrix}$$

Calculați entropia în biți, entropia maximă și entropia relativă.

Se dă $\log_2 5 = 2,3219$.

6. Bifați afirmațiile GREȘITE cu privire la distribuția normală:
- a. distribuția normală are forma unui clopot cu două cozi infinite ce tind asimptotic la zero.
- b. există o singură distribuție normală.
- c. distribuția normală se mai numește și distribuție laplaceană.
- d. distribuția normală este caracterizată de medie și dispersie.
7. În urma aplicării unui test binar pentru diagnosticarea unei boli s-au obținut următoarele rezultate:

TEST	+	-
NX	8	203
X	60	11

Calculați sensibilitatea, specificitatea, *VPP*, *VPN* și prevalența bolii.

8. Se dă următorul șir bidimensional reprezentând un eșantion aleatoriu de 4 perechi de valori:

x:	2	3	3	4
y:	2	2	3	4

Știind că în fenomen există o legătură între cele două variabile

- a. să se reprezinte șirul într-o diagramă de împrăștiere
- b. să se calculeze coeficientul de corelație liniară
- c. să se calculeze dreapta de regresie a lui y în x