

Koduse kontrolltöö juhend

Teadmiseks

- Kahe märgatavalt sarnase kodutöö esitamisel erinevate tudengite poolt lähevad töö eest antavad punktid esitajate vahel “Bingo Loto Jackpot’i” põhimõttel jagamisele.
- Koduse kontrolltöö esitamise tähtaeg on **9. juuni 2014 kell 09.00**. Kodutöö tuleb saata hiljemalt määratud tähtajaks aadressil kristiina@itcollege.ee. **Hilinenud tööd arvesse ei lähe.**
- Kodutöö väärtus on 15 punkti.
- Töö tuleb vormistada *MS Exceli* failina, mille nimi **peab olema** kujul *eesnimi_perekonnanimi.xlsx* (või *eesnimi_perekonnanimi.xls*).

Iseseisva töö aluseks on reaalne andmestik, mida töö teostaja peab oskama täpselt kirjeldada (mida kajastab, mis ajast, kust pärit jne). Juhul kui vaadeldavad tunnused sõltuvad ajast (kaubaartikli hind, sündimus, suremus, palk, liiklusõnnetuste arv) peavad olema kaasatud andmed viimase hetkel kättesaadava ajaperioodi kohta.

Andmestik peab sisaldama 20-30 objekti andmeid, nende hulgas peab olema vähemalt 2 arvtunnust. Andmestiku võib otsida huvipakkuva ala kohta Internetist (või raamatukogust) või ka näiteks Eesti Statistikaameti väljaannetest (Internetist, samuti raamatukogust). Andmestikuks pole lubatud valida loengus või praktikumis vaadeldud andmetabeleid ja õpikutes näiteülesannetena kasutatud materjale.

Andmetabeli kahe valitud tunnuse väärtuspaaride põhjal tuleb **loengukursuses esitatud valemite abil** (programmis *MS Excel*, mitte kasutada teisi statistilise andmetöötluse süsteeme) leida järgmised arvnäitajad ja (või) tabelid ning konstrueerida illustreerivad joonised. Saadud tulemusi tuleb kontrollida (kui võimalik) *MS Exceli* statistikafunktsioonide ja protseduuride abil.

1. Andmestiku kirjeldus (kust pärit, mida kajastab, mis ajast). Lisada viide Internetti või raamatule.
2. Leida Pearsoni empiiriline korrelatsioonikordaja r . Selgitada leitud korrelatsioonikordaja sisulist tähendust (selgitada tuleb nii arvu märki kui ka arvu absoluutväärtust).
3. Kontrollida korrelatsiooni olemasolu olulisusnivooga 5%. Sõnastada järeldus.
4. Määrata sõltuv muutuja (tagajärg) ja sõltumatu muutuja (põhjus). Põhjendada oma valikut.
5. Arvutada regressioonisirge parameetrid a ja b . Kirjutada välja regressioonisirge võrrand.
6. Arvutada determinatsioonikordaja r^2 . Selgitada näitaja sisulist tähendust (selgitada tuleb kui hästi mudel töötab ja mitu protsenti sõltuva muutuja kogumuudust on selgitatav mudeliga).
7. Joonestada graafik, millele on kantud antud punktid (x_i, y_i) , regressioonisirge $\hat{y} = a + bx$ ja determinatsioonikordaja.
8. Leida regressioonisirge parameetrite a ja b 90%-lised usalduspiirkonnad.
9. Leida statistilise prognoosi punkthinnang \hat{y}_p ning vastav 90%-line usalduspiirkond mingis andmestikuga kaetud punktis. Kommenteerida tulemust.
10. Kasutatud materjalide loetelu.

Kõik saadud tabelid, arväärtused ja joonised peavad olema selgitatud ja kommenteeritud. Lisaks peab tudeng olema võimeline selgitama ja kommenteerima oma kodutöö teksti ning failis *eesnimi_perekonnanimi.xlsx* (või *eesnimi_perekonnanimi.xls*) sisalduvat analüüsi.

Abimaterjalina töö teostamisel võib kasutada näiteks järgnevaid materjale:

- ✓ Õppematerjalid ja vastavad lisad näiteülesannete lahendustega Moodles.
- ✓ H. Käerdi. Statistika ja tõenäosusteooria alused. Tallinn, 1999.
- ✓ A. Kiviste. Matemaatiline statistika MS Excel keskkonnas. Tallinn, 1999.