Probabilitat i Estadística 1
Grau de Ciència i Enginyeria de Dades
Problemes Tema 2

Variables aleatòries i models discrets

1. Considerem una successió de cinc tirades d'una moneda perfecta. Sigui X el nombre de vegades que una cara és seguida immediatament d'una creu. És a dir, que si $\omega = C + C + C$, aleshores $X(\omega) = 2$. Trobeu les probabilitats $P(X = k)$.
   Indicació: *Calculeu primer* $P(X = 2)$ *i* $P(X = 0)$.

   Resultat: $P(X = 0)\frac{3}{16}, P(X = 1) = \frac{10}{16}, P(X = 2) = \frac{3}{16}$.

2. Sigui X una variable aleatòria que té per funció de distribució:

$$F(x) = \begin{cases} 0 & \text{si } x < -1 \\ \frac{1+x}{9} & \text{si } -1 \leq x < 0 \\ \frac{2+x^2}{9} & \text{si } 0 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

   Calculeu:

   (a) $P(X = 2)$
   (b) $P(X \in [-\frac{1}{2}, 3))$
   (c) $P(X \in \{(-1, 0] \cup (1, 2)\})$
   (d) $P(X \in \{[0, \frac{1}{2}) \cup (1, 2]\})$

   Resultat: (a) 1/3; (b) 17/18; (c) 5/9; (d) 29/36.

3. Suposem que una urna conté 7 boles vermelles i 3 de blaves. Si se seleccionen 5 boles aleatòriament, sense reemplaçament, determinar la funció de probabilitat del nombre de boles vermelles que s'obtenen.

   Resultat: $P(X = k) = \frac{\binom{7}{k}\binom{3}{5-k}}{\binom{10}{5}}$ amb $k = 2, 3, 4, 5$.

4. Es llença 9 vegades una moneda trucada amb probabilitat de cara 0.6. Calcular la probabilitat d'obtenir un nombre parell de cares.

   Resultat: 0.4999997.

5. Una màquina de fabricar ampolles té una probabilitat igual a 0.1 de produir una ampolla defectuosa. Suposem que cada dia fabrica deu ampolles. Aquestes ampolles passen un control de qualitat i les defectuoses es retiren. Sigui $p_1 = 0.2$ la probabilitat que una ampolla defectuosa sigui classificada malament i $p_2 = 0.05$ la probabilitat que una ampolla no defectuosa sigui classificada malament. Sigui X el nombre d'ampolles classificades com a defectuoses.

   (a) Obteniu una expressió per a $\Pr(\{X = k\})$.
   (b) Calculeu $\Pr(\{X = 3\})$ i $\Pr(\{X > 3\})$.
   (c) Sigui Y el nombre d'ampolles defectuoses entre les que s'han classificat com a defectuoses. Quina distribució condicionada té Y donat que $X = x$?

   Resultat: (a) $\Pr(X = k) = \binom{10}{k} p^k (1 - p)^{10-k}$; (b) 0.0920 i 0.0274; (c) $B(x, 0.64)$.

6. Un vigilant ha d'obrir una porta durant la nit. Té un clauer amb deu claus diferents i escull una clau a l'atzar seguint un dels mètodes següents:

   A: Prova amb cada una de les claus, sense repetir-ne cap.

   B: Prova una clau, la barreja amb les altres i en prova una altra.

Sigui X_A (respectivament X_B) la variable aleatòria que designa el nombre de claus assajades amb el procediment A (respectivament B).

(a) Determineu les distribucions de X_A i X_B.

(b) Calculeu la probabilitat d'assajar més de vuit claus en cadascun dels dos casos.

(c) Cada vespre el vigilant tira una moneda que té probabilitat de cara 2/3. Si surt cara utilitza el mètode A per obrir la porta i si surt creu utilitza l'altre mètode. Calculeu la probabilitat que hagi sortit creu si sabem que amb vuit intents no ha pogut obrir la porta.

Resultat: (a) $X_A \sim U(10)$ i $X_B \sim Geom(1/10)$; (b) 0.2 i 0.43; (c) 0.52.

7. Entre dues-centes famílies amb quatre fills cadascuna, quantes cal esperar que tinguin:

   (a) Tres nens i una nena,

   (b) Dos nens i dues nenes,

   (c) Dues o tres nenes.

   Suposeu que les probabilitats de tenir un nen o una nena són iguals.

   Resultat: (a) 50; (b) 75; (c) 125.

8. **(Machine Learning)**. Es disposa d'un conjunt de classificadors automàtics que classifiquen nous exemples en una de dues categories possibles, només una de les quals és la categoria a la qual realment pertany cada exemple. Podem pensar que els *exemples* són clients que demanen un préstec a un banc, que els vol classificar com pagadors *bons* o *dolents*. Els classificadors automàtics podrien ser diferents qüestionaris tipus test que els clients han d'omplir.

   Per tal de millorar la precisió de les classificacions individuals, es combinen 3 d'aquests classificadors: quan arriba un nou exemple, és classificat per cadascun dels 3 classificadors i la decisió de classificació conjunta es pren per majoria. Si suposem que cada classificador individual té una probabilitat $p = 0.7$ de classificar correctament un nou cas i que actuen de forma independent, calcular la probabilitat que el nou cas es classifiqui correctament quan els combinem.

   Resultat: 0.784.

9. Designem per X el nombre de cotxes que arriben al peatge d'una autopista durant un minut. Suposem que X segueix una llei de Poisson de paràmetre 5. Calculeu:

   (a) Probabilitat que passin 5 cotxes entre les $3^h\,4^m$ i les $3^h\,5^m$.

   (b) Probabilitat que passin més de 2 cotxes entre les $3^h\,4^m$ i les $3^h\,5^m$.

   (c) Probabilitat que passin més de 6 cotxes entre les $3^h\,4^m$ i les $3^h\,6^m$. *(Suposeu que el nombre de cotxes que passen en dos minuts diferents són independents.)*
   Indicació: Podeu fer servir aquest resultat: *Si $X_1 \sim Poisson(\lambda_1)$ i $X_2 \sim Poisson(\lambda_2)$ són dues variables aleatòries independents, aleshores $X = X_1 + X_2 \sim Poisson(\lambda_1 + \lambda_2)$.*

   (d) S'instal·la al peatge un comptador per comptar els cotxes que passen. Aquest té una probabilitat 0.1 de fallar no comptant un cotxe que passa. Trobeu la distribució del nombre de cotxes per minut que dóna el comptador.

   Resultat: (a) 0.175; (b) 0.8753; (c) 0.8698; (d) Poisson (0.9λ).

10. El centre de càlcul d'una important empresa atén les incidències que els sorgeixen als treballadors. S'ha observat que les incidències apareixen esporàdicament, encara que l'elevat nombre d'usuaris implica que el volum de problemes a tractar diàriament sigui considerable (s'ha suposat una mitjana de 2.35 incidències/dia). Es suposa que les incidències de dies diferents són independents.

    (a) Probabilitat que en un dia es produeixi 3 incidències.

    (b) Probabilitat d'observar menys de 3 incidències en un dia.

    (c) Entre el dilluns i el dimarts han rebut sis incidències. Quina és la probabilitat que de dilluns a divendres es tractin no més de 15?

(d) Quina és la probabilitat que cap dia de la setmana laboral presenti incidències?

Indicació: Podeu fer servir aquest resultat: *Si* $X_1 \sim Poisson(\lambda_1)$ *i* $X_2 \sim Poisson(\lambda_2)$ *són dues variables aleatòries independents, aleshores* $X = X_1 + X_2 \sim Poisson(\lambda_1 + \lambda_2)$.

Resultat: (a) 0.2063; (b) 0.5828; (c) 0.8254; (d) 0.0000079.

11. Una companyia aèria ven 200 bitllets per un vol en un avió que només té 198 places perquè sap que, en mitjana, un 1% dels viatgers no es presenten a la sortida dels vols. Determinar la probabilitat que tots els viatgers que es presentin a la sortida d'aquest vol tinguin seient.

Resultat: 0.594.

12. Es diu que una variable aleatòria X amb valors en $\mathbb{N}$ no té memòria si per a tot $n \in \mathbb{N}$, $P\{X > n\} > 0$, i per a qualssevol n, $k \in \mathbb{N}$, $\Pr(\{X > k + n | X > k\}) = P\{X > n\}$. Proveu que X no té memòria si i només si X és geomètrica.

13. Considereu un sistema informàtic al qual els treballs arriben seguint un procés de Poisson amb una mitjana de 2 treballs per minut. Determineu la probabilitat que en qualsevol interval d'un minut hi hagi ...

    (a) 0 treballs,
    (b) exactament 2 treballs,
    (c) com a molt 3 arribades.
    (d) Quin és el màxim de treballs que ha d'arribar en un minut amb 90 % de certesa?

Resultat: (a) 0.135; (b) 0.271; (c) 0.857; (d) 4.

14. Hem començat una col·lecció de cromos que consta de n cromos diferents. Comprem els cromos d'un a un i el venedor ens assegura que la probabilitat d'obtenir un cromo de qualsevol tipus és la mateixa. Quants cromos esperem comprar per a completar la col·lecció? Calculeu aquest valor per $n = 10$.

Resultat: 29.29.

15. Suposem que dues loteries tenen n bitllets cada una i idèntic premi. En termes del guany esperat, és millor comprar dos bitllets de la mateixa loteria o comprar un de cada loteria?

16. **(Paradoxa de Sant Petersburg)** Es juga a cara o creu amb una moneda equilibrada. Suposem que sempre escollim cara. Comencem apostant 1 euro i doblem l'aposta cada cop que perdem. Quan guanyem un cop, abandonem el joc. Proveu que amb aquesta estratègia sempre es guanya 1 euro. Quin és el valor esperat del nombre d'euros apostats?

Resultat: ∞.

Variables contínues i models continus

17. Un procés de fabricació produeix fils de longitud aleatòria. La longitud X segueix una distribució de probabilitat contínua amb funció de densitat

$$f(x) = \begin{cases} kxe^{-x} & \text{si } x \geq 0, \\ 0 & \text{altrament.} \end{cases}$$

    (a) Calculeu el valor de la constant k i doneu l'expressió per a la funció de distribució.
    (b) Calculeu $\Pr(1 < X < 3), \Pr(X > 5)$ i $\Pr(3.9 < X < 4.1)$.

Resultat: (a) 1, $F_X(x) = 1 - (x + 1)e^{-x}\mathbf{1}_{[0,\infty)}$; (b) 0.5366, 0.0404 i 0.014665.

18. Siguin f i g dues funcions de densitat. Demostreu que $\alpha f + (1 - \alpha)g$ és una densitat, a on $0 \leq \alpha \leq 1$.

19. Suposeu que la v.a. X té funció de densitat f. Definim $Y = aX + b$ ($a \neq 0$). Determineu la funció de densitat de Y.

20. La *"vida útil d'un transistor mesurada en anys"*, X, segueix una distribució contínua amb funció de densitat que decreix exponencialment:

$$f(x) = \begin{cases} ke^{-x} & \text{si } x \geq 0, \\ 0 & \text{altrament.} \end{cases}$$

Calculeu el valor de k i el valor esperat de X.

Resultat: 1 i 1.

21. Sigui

$$f(x) = \begin{cases} \frac{1+\alpha x}{2} & \text{si } -1 \leq x \leq 1 \\ 0 & \text{altrament,} \end{cases}$$

on $-1 \leq \alpha \leq 1$.

(a) Demostreu que f és una densitat i trobeu la funció de distribució corresponent.

(b) Calculeu $E(X)$ i $Var(X)$.

(c) Calculeu la mediana de X.

Resultat: (b) $\frac{1}{3}\alpha$ i $\frac{3-\alpha^2}{9}$; (c) $\frac{-1}{\alpha} \pm \frac{\sqrt{1+\alpha^2}}{\alpha}$, segons el signe d'$\alpha$.

22. Suposeu que X és una v.a. contínua amb funció de densitat $f(x) = (1/2)e^{-|x|}$ per $-\infty < x < +\infty$. Determineu el valor $x_{0.9}$ tal que $F(x_{0.9}) = 0.9$, essent F la funció de distribució de X.

Resultat: 1.609.

23. Suposem que la durada en minuts d'una trucada a llarga distància segueix una llei exponencial de paràmetre 1/5. Trobar la probabilitat que la durada d'una conversa:

(a) superi els cinc minuts;

(b) estigui entre cinc i sis minuts;

(c) sigui inferior a tres minuts;

(d) sigui inferior a sis minuts sabent que es superior a tres minuts.

Resultat: (a) 0.3678; (b) 0.0666; (c) 0.4512; (d) 0.4512.

24. La duració en hores, T, d'un tub electrònic segueix una distribució exponencial de mitjana μ hores. A un dels controls de qualitat mostregem un lot, fem funcionar (de forma accelerada) aquests tubs durant 30 hores i seleccionem com a bons aquells tubs que no s'han espatllat durant la prova.

(a) Calculeu la funció de distribució dels tubs bons.

(b) Quina és la duració esperada dels tubs bons?.

Resultat: (a) $1 - e^{-\frac{t-30}{\mu}}$; (b) $30 + \mu$.

25. Trobeu la mediana, el primer i el tercer quartil d'una distribució exponencial de paràmetre β.

Resultat: $\ln 2/\beta, -\ln(3/4)/\beta$ i $\ln 4/\beta$.

26. Si T és una v.a. exponencial de paràmetre β tal que $\mathrm{P}(T < 1) = 0.05$, quin és el valor de β?.

Resultat: 0.0513.

27. Suposeu que un cert sistema conté tres components que funcionen independentment una de l'altra. Suposeu que la vida de la primera component, mesurada en hores, segueix una distribució exponencial amb paràmetre $\beta = 0.001$; la vida de la segona component segueix una distribució exponencial amb paràmetre $\beta = 0.003$; i la vida de la tercera component segueix una distribució exponencial amb paràmetre $\beta = 0.006$. Determineu la probabilitat que el sistema funcioni al cap de 100 hores en les dues situacions següents:

(a) Les tres components estan connectades en sèrie.

(b) Les tres components estan connectades en paral·lel.

Resultat: (a) 0.3678; (b) 0.988.

28. Proveu que la següent instrucció d'R genera valors d'una exponencial amb esperança igual a 1000:

```
x <- -1000*log(runif(1))
```

29. Suposem que la mesura del voltatge d'un circuit elèctric segueix una distribució normal de mitjana 120 i desviació estàndar 2. Es fan tres mesuraments independents d'aquest voltatge. Quina és la probabilitat que els tres mesuraments estiguin entre 116 i 118?

Resultat: 0.0025.

30. Suposeu que el coeficient d'intel·ligència (CI) dels alumnes universitaris es distribueix normalment amb mitjana 120 i variància 150.

    (a) Si una classe de 50 alumnes es seleccionada a l'atzar, quina és la probabilitat que el coeficient d'intel·ligència mig sigui superior a 125?

    (b) Si seleccionem dos grups de 50 estudiants a l'atzar, quina és la probabilitat que el coeficient mig de les dues classes difereixi en més de 5 punts?

    (c) Necessitem 5 estudiants per formar un equip per participar en una olimpíada universitària. Anem entrevistant estudiants fins a obtenir 5 amb CI més gran que 125. Calculeu la probabilitat d'haver d'entrevistat 11 estudiants.

Resultat: (a) 0.001946.; (b) 0.412269; (c) 0.07926.

31. La longitud dels peus del homes adults d'un determinat país es distribueix segons una variable aleatòria normal amb esperança 24.5 cm. i desviació típica 1 cm. Un fabricant de sabates treu al mercat un model del qual presenta quatre mides. El de mida més petita és adequat per peus de longituds compreses entre 23 i 24 cm., i els següents cobreixen els peus de longituds entre 24 i 25 cm, entre 25 i 26 cm, i entre 26 i 27 cm., respectivament.

    (a) Quina proporció d'homes trobaran el seu número entre els fabricats d'aquest model?

    (b) Si el fabricant decideix afegir un número més a la seva col·lecció, què hauria de fer, fabricar-ne un de més petit o un de més gran?

Resultat: (a) 0.927; (b) el petit.

32. Els litres de benzina distribuïts cada dia per una benzinera és una variable aleatòria normal de mitjana 15.000 litres i desviació típica 1.000 litres. La benzinera compra el litre de benzina a 0,5 euros y el ven a 0,8 euros.

    (a) Quina quantitat diària s'ha de tenir preparada per tal de poder satisfer la demanda el 95% dels dies?

    (b) Quin percentatge de dies els beneficis superen els 5.000 euros?

Resultat: Amb la taula de la $N(0,1)$: (a) 16645; (b) 0.0475.

33. En una ferreteria es venen capses de claus. El nombre de claus de cada capsa segueix una distribució normal de paràmetres $\mu = 200$ i $\sigma = 10$.

    (a) Quin percentatge de capses contenen entre 180 i 220 claus?

    (b) Si es torna l'import de les capses que contenen menys de 180 claus i comprem dues capses, amb quina probabilitat es tornaran l'import de les dues capses?

Resultat: (a) 0.9544; (b) 0.00052.

34. L'error de mesura X del temps d'un procediment és normal amb esperança 0 i desviació estàndard $\sigma = 1/4$ segons. Es considera que les mesures es repeteixen de forma independent.

    (a) Probabilitat que l'error sigui menor de 0.1 segons.

    (b) Error màxim (amb probabilitat 0.95) en una mesura.

    (c) Error màxim (amb probabilitat 0.95) en la mitjana de 10 mesures.

(d) Nombre n mínim de mesures per tal que l'error màxim de la mitjana de les n mesures (amb probabilitat 0.95) sigui inferior a 0.1 segons.

Resultat: (a) 0.3108; (b) 0.49 segons; (c) 0.155 segons; (d) 25.

35. Si $X \sim N(\mu, \sigma^2)$. Calculeu el valor dels seus quartils i del rang interquartílic.