

NOM: _____ COGNOMS: _____
(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs)

Problema 1 (B1-B2)

Tenim una moneda trucada de tal manera que la probabilitat de cara és $\frac{1}{4}$ i la de creu és $\frac{3}{4}$, i la llencem 4 vegades seguides.

1.- (1 punt) Indiqueu l'arbre de l'experiència aleatòria, i el conjunt de resultats amb les seves probabilitats.

Resultats:

2.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que surtin primer 2 cares seguides i després 2 creus. I que surtin primer 2 creus seguides i després 2 cares.

3.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que surtin 3 cares i una creu

4.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que a la quarta tirada surti cara quan a la primera ja ha sortit cara

5.- (1 punt) Calculeu i justifiqueu formalment quina és la probabilitat que surtin més cares que creus

Seguint amb la moneda anterior que hem llençat 4 vegades seguides:

6.- (2 punts) Definiu les variables aleatòries nombre de cares i nombre de creus, indicant la funció de probabilitat i la funció de distribució de probabilitat i la seva representació gràfica

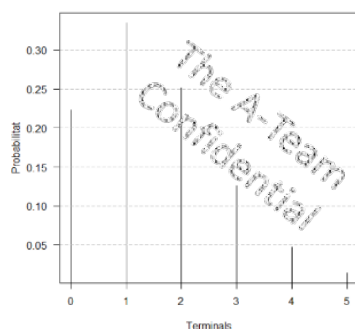
7.- (1 punt) Calculeu l'esperança i la variància de les variables aleatòries anteriors

8.- (2 punts) Indiqueu i justifiqueu la taula de probabilitat conjunta de les dues variables anteriors. Calculeu la seva covariància i correlació i relacioneu els resultats amb el fet de ser dependents o independents

Problema 2 (B3-B4)

Els sistemes de detecció de possibles atacs informàtics de la UPC han previst un assalt imminent d'un equip d'hackers (*Equip X*) als sistemes de la Universitat. Només un equip format per 4 ex-alumnes de la FIB (*Equip A*) té els coneixements suficients per evitar aquest atac. Malgrat que, actualment, es troben en parador desconegut, la UPC els ha aconseguit trobar i els contractarà per aquesta tasca. **(Totes les qüestions valen 1 punt excepte les 2 últimes que valen 2 punts cadascuna)**

1. Hi ha 10 servidors a la UPC que proporcionen servei a les unitats fonamentals. Si fallen més de 5 d'aquests servidors, la UPC haurà de tancar. L'Equip A creu que la probabilitat de que el Equip X deixi inoperatius cadascun d'aquests servidors és de 0.6. Quina és la probabilitat que la UPC hagi de tancar?
2. L'Equip A creu que pot introduir una millora en els *Firewalls* dels servidors per a que la probabilitat de que tanqui la UPC sigui de 0.102. Quina és ara la probabilitat de que quedi inoperatiu cadascun dels servidors?
3. Entre els informes de l'Equip A, tenen el gràfic de la funció de probabilitat del número de terminals infectats per segon per un virus informàtic en els atacs de l'Equip X i saben que es distribueix de forma poissoniana. Digues a partir d'aquesta informació, quin és el temps esperat entre les infeccions consecutives de dos terminals.



4. L'Equip A està capacitat per desactivar l'atac de l'equip X, un cop es comenci, en 3 minuts. Quina és la probabilitat que en aquest temps NO s'hagin infectat més de 250 terminals de la universitat?
5. La UPC ha demanat un pressupost a l'Equip A per dur a terme la missió en funció del número d'hores invertides. Com que aquest número d'hores és variable i desconegut a priori, la UPC es conforma amb que se li proporcioni un interval dels euros pels qual es tingui una probabilitat de 0.90 de que l'import final estigui dins de l'interval. Es sap que la distribució de les hores invertides en missions similars ha estat $N(\mu=40, \sigma=10)$ i que el preu/hora és de 50€. Proporciona aquest interval.

La UPC vol tenir informació a priori de la magnitud dels atacs informàtics perpetrats per l'Equip X. Per aquest motiu, ha demanat a l'Equip A que li passi algunes dades sobre el percentatge (**P**) d'ordinadors afectats resultat d'atacs previs en altres institucions. L'Equip A proporciona la següent informació sobre aquest percentatge d'ordinadors afectats en els darrers 30 atacs pels quals disposen informació

Report Equip X		18 de juny de 2021										
$\sum_{i=1}^{30} P_i = 1575.7$ $\sum_{i=1}^{30} P_i^2 = 85609.77$	<i>The A-Team Confidential</i>	<i>Taula 1. Quantils del percentatge d'afectació en els darrers atacs perpetrats per l'equip X.</i>										
		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
		35.8%	42.7%	44.3%	46.5%	50.0%	51.2%	53.5%	55.8%	60.3%	63.0%	76.9%

6. Calcula una estimació puntual de la mitjana d'aquest percentatge i del seu error estàndard.

7. Fes un contrast d'hipòtesi per veure si la mitjana poblacional és del 50% o no amb un $\alpha=5\%$. 1) Planteja el test; 2) Digues les premisses; 3) Calcula l'estadístic; i 4) Conclou sobre la prova

8. Fes un contrast d'hipòtesi per veure si la proporció poblacional d'atacs amb més del 50% d'ordinadors afectats és de 0.5 o superior amb un $\alpha=1\%$. 1) Planteja el test; 2) Digues les premisses; 3) Calcula l'estadístic; i 4) Conclou sobre la prova

NOM: _____ COGNOMS: _____
(Contesteu cada pregunta en el seu lloc. Expliqueu i justifiqueu els càlculs)

Problema 3 (B5-B6)

Per tal de mesurar distàncies en xarxes de comunicacions de forma indirecta, uns investigadors han publicat un treball on descriuen que utilitzen el nombre de *hops* (nombre de salts entre nodes) emprats per els missatges viatjant entre clients i servidors. Els investigadors han distingit entre xarxes als Estats Units (US) i resta del món (no-US), segons la localització dels extrems. Llavors, tenen 364 observacions a US i 417 a no-US: la mitjana dels *hops* a US és de 18.64 (20.16 a no-US), i la desviació tipus és 13.6 (14.4 a no-US)

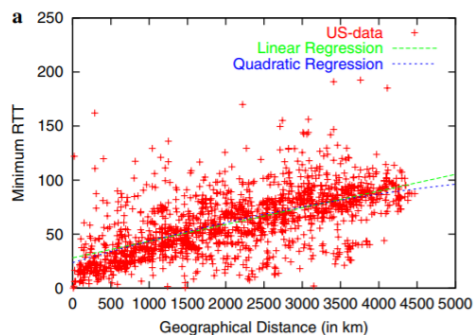
1.- (1.5 punts) Desenvolpeu una prova d'hipòtesis formal per demostrar si hi ha diferències en el nombre mitjà de *hops* entre ambdues poblacions. Recordeu definir correctament les hipòtesis, premisses, estadístic de la prova amb la seva distribució de probabilitat, zones d'acceptació i rebuig per a un risc del 5%, resolució i interpretació de la conclusió.

2.- (1 punt) Trobeu també el p-valor del resultat anterior i expliqueu formalment què significa (*no feu una interpretació pràctica, del tipus "si és major que ... llavors ..."*)

3.- (1 punt) Amb quins arguments es podria haver plantejat una prova unilateral, i amb quin objectiu?

4.- (1.5 punts) En les dades de US s'han comptat 23 casos en els que el missatge ha estat rebutjat i va haver de ser reenviat de nou. En l'altre conjunt, això va ocórrer 40 vegades. Es pot dir que el rebuig és més freqüent a les xarxes de la resta del món? Estructureu la resposta d'acord al procediment de la prova d'hipòtesis.

El treball d'aquells investigadors se centra en la relació entre la distància geogràfica entre origen i destí, i RTT (round trip time, ms), per tal de dirigir a l'usuari al servidor amb la menor latència. La correlació entre les dues variables a les dades US s'indica a l'article (0,6113), però no l'equació de la recta de regressió. No obstant això, hem pogut determinar que el centre del núvol de punts està en (2287km, 62.677ms). D'altra banda, visualment s'ha estimat que les desviacions respectives serien aproximadament 1300km i 33ms



5.- (1.5 punts) Amb aquesta informació, calculeu l'equació de la recta de regressió i estimi el temps RTT que trigaria una connexió entre dos punts separats 1500km.

6.- (1 punt) Calculeu també el valor de la desviació residual i el coeficient de determinació, explicant quina informació dóna cada un dels dos indicadors.

7.- (1.5 punts) Interpreteu el valor del pendent i obtingui un interval de confiança a l'95% per al pendent poblacional.

8.- (1 punt) Indiqueu les premisses del model lineal, què impliquen i com es comproven. A la vista del gràfic indicat plantegeu per cadascuna de les premisses si seria assumible o no i per què.