

Indicacions

La parametrització de la distribució Binomial Negativa que heu de fer servir és:

$$X \sim NB(r, p) \quad r > 0, p \in [0, 1] \quad X \in \{0, 1, 2, \dots\} \quad P(X = x) = \binom{r+x-1}{x} p^r (1-p)^x$$

$$E(X) = \mu = \frac{r(1-p)}{p}$$

$$V(X) = \sigma^2 = \frac{r(1-p)}{p^2}$$

(El paràmetre r pot prendre qualsevol valor real positiu)

Recordeu que $\binom{r+x-1}{x} = \frac{\Gamma(r+x)}{\Gamma(r)\Gamma(x)}$

En l'apartat A, feu servir els dos primers moments no centrats per aplicar el principi de substitució.

En l'apartat B heu de calcular el màxim de la funció de log-versemblança. Es pot fer resolent les equacions de versemblança, o bé, calculant el màxim directament. En tots els casos, en part s'ha de resoldre numèricament.

Algunes funcions que en aquest entregable podeu necessitar, en la notació del R són:

- `gamma(x)`: la funció $\Gamma(x)$, `lgamma` si volem el seu logaritme.
- `factorial(x)`: per obtenir $x!$, `lfactorial` si volem el seu logaritme.
- `psigamma(x)`: és la funció $\Psi(x) = (\log \Gamma(x))'$.

Algunes comandes del R que us poden ser útils són:

- `uniroot`: per resoldre una equació.
- `optimize`: per trobar el mínim, o el màxim, d'una funció univariant.
- `nlm`: per trobar el mínim d'una funció multivariant, per exemple -funció de log-versemblança.
- `optim`: com l'anterior, per trobar el mínim (o el màxim) d'una funció multivariant.

Una cosa que heu de tenir en compte és la precisió, en el qüestionari us demanem 5 xifres significants, en totes les comandes del R hi opcions que us poden servir per augmentar la precisió, com per exemple:

- Modificar la tolerància, augmentar el nombre d'iteracions.
- Donar els valors inicials, o l'interval, més ajustats.
- Donar el gradient de la funció a optimitzar.

L'interval de predicció amb probabilitat *prob* que es demana en la qüestió 11 fa referència a l'interval $[a, b]$ tal que: $P(X < a) = P(X > b) \leq \frac{1-prob}{2}$ i per tant $P(a \leq X \leq b) \geq prob$. Tingueu en compte que la variable és discreta i que com els valors de a i b pertanyen al suport de la variable, són valors sencers. Per això no es pot garantir que la probabilitat entre els dos valors sigui exactament *prob*. La pregunta 12 fa referència al càlcul de la probabilitat correcta de pertànyer a l'interval calculat, la qual sabem que ha de ser igual o superior a *prob*.

Indicacions per l'apartat C: La informació que teniu de la distribució Beta, a la taula de distribucions habituals, us pot ser de gran utilitat.

Pel càlcul de la esperança i de la variància de X a posteriori feu servir la llei de la esperança iterada.