Pràctica 1. Població mundial

Equacions Diferencials i Modelització I

4 de octubre de 2022

1 Introducció

La taula següent[1] mostra la població mundial, $p(t)$, en milers de milions entre els anys 1950 i 2005:

t	$p(t)$	t	$p(t)$
1950	2.535093	1980	4.451470
1955	2.770753	1985	4.855264
1960	3.031931	1990	5.294879
1965	3.342771	1995	5.719045
1970	3.698676	2000	6.124123
1975	4.076080	2005	6.514751

El problema que ens plantegem en aquesta pràctica és donar un model teòric que permeti trobar la funció $p(t)$ per tal de poder fer prediccions de la població del món en els propers anys. Al llarg de la pràctica veurem algunes aproximacions a aquest problema que dependran del model escollit i, per tant, les respostes a preguntes del tipus "Quina és la població del món l'any 2050?", dependran del model triat.

Representa gràficament, $(t, p(t))$, aquestes dades. Quina funció diries que té aquesta gràfica?

```
Xt=[1950,1955 .. 2005];Xt
Yt=[2.535093,2.770753,3.031931,3.342771,3.698676,4.076080,4.451470,
4.855264,5.294879, 5.719045,6.124123,6.514751];
i=var(``i'');taula=points([[Xt[i],Yt[i]] for i in
range(0,len(Xt))]); show(taula);
```

També es pot obtenir la gràfica mitjançant

```
list_plot([[Xt[i],Yt[i]] for i in range(0,len(Xt))])
```

2 Presentació dels models

Per definició, la **taxa de creixement demogràfic**, TC, d'un determinat territori (país, regió, província, ciutat, etc.) és l'augment de la població durant un període de temps fixat respecte la població a l'inici del període.

[1]Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, World Population Prospects: The 2006 Revision and World Urbanization Prospects: The 2005 Revision, `http://www.un.org/esa/population/publications/WUP2005/2005wup.htm`

Si es pren un model continu es defineix TC com p'/p. En aquesta pràctica ens centrarem en els models continus així que si denotem per $f(p)$ la funció que defineix TC, els models de població que estudiarem seran de la forma:

$$p' = pf(p), \quad p(t_0) = p_0.$$

Omple la taula següent usant l'aproximació $p'(t) = (p(t+h) - p(t))/h$:

t	$p(t)$	$p'(t)$	$p'(t)/p(t)$	t	$p(t)$	$p'(t)$	$p'(t)/p(t)$
1950	2.535093			1980	4.451470		
1955	2.770753			1985	4.855264		
1960	3.031931			1990	5.294879		
1965	3.342771			1995	5.719045		
1970	3.698676			2000	6.124123		
1975	4.076080			2005	6.514751		

```
dY=[(Yt[i+1]-Yt[i])/(Xt[i+1]-Xt[i]) for i in range(0,len(Xt)-1)];
dYY=[dY[i]/Yt[i] for i in range(0,len(dY))];
```

Fes la gràfica de $(p, p'/p)$ i digues quina és la funció que prendries com a taxa de creixement.

```
taula0=points([[Yt[i],dYY[i]] for i in range(0,len(dY))]);
show(taula0);
```

3 Model de Malthus

Thomas Malthus l'any 1798, al seu "Assaig sobre el principi de la població", va proposar el primer model de creixement, on assumia una TC constant: $f(p) = a$.

Troba el valor de la mitjana dels valors de p'/p calculats en la taula anterior.

```
a0=mean(dYY);
```

Dóna la funció solució del model de Malthus, $p(t)$, que faries servir per a predir la població del món, tot resolent l'equació

$$p' = ap, \; p(1950) = 2.535093.$$

```
t=var('t');p=function(``p'')(t);
odesol=desolve(diff(p(t),t)-a0*p(t),[p,t],[Xt[0],Yt[0]])
```

Remarquem que 'desolve' dóna com a resultat una funció (i no una expressió)!

Fes la gràfica dels valors de la taula juntament amb la solució que has trobat.

```
plot(odesol(t),(t,1950,2005));
```

Quin serà el valor de la població l'any 2050, $p(2050)$?

```
n(odesol(2050))
```

4 Model de Verhulst (I)

Pierre François Verhulst, l'any 1838, al treball "Notice sur la loi que la population poursuit dans son accroissement. Correspondance mathématique et physique 10:113–121." va proposar un nou model que tenia en compte, entre d'altres, la competència per l'aliment, la mortalitat,... Aquest model conegut com model logístic respon a una taxa de creixement lineal

$$f(p) = ap + b.$$

Seguint aquest model, busca la recta de regressió que més s'ajusta al conjunt de punts $(p, p'/p)$.

```
Yp=[Yt[i] for i in range(0,len(dY))]; YpdYY=[[Yp[i],dYY[i]] for i in
range(0,len(dY))]; var('a,b'); model(p)=a*p+b;
recta0=find_fit(YpdYY,model); recta0;
```

Quin és el coeficient de correlació d'aquests punts? (Aquest coeficient mesura si per les nostres dades $(p, p'/p)$ es vegi una relació lineal, i així la regressió lineal tingui sentit. La seva definició podeu trobar a la secció 7).

Quina és la suma dels quadrats de les diferències en aquest cas? (És una mesura per l'error emprat usant la regressió lineal: $\sum_{i=0}^{10}(p_i'/p_i - (ap_i + b))^2$)

Quina era en el cas de prendre $f(p) = a_0$ constant? (És a dir: $\sum_{i=0}^{10}(p_i'/p_i - a_0)^2$)

Representa, en la mateixa gràfica, els punts i la recta de regressió.

```
points(YpdYY)+plot(model(a=recta0[0].rhs(),b=recta0[1].rhs()),(p,2,6.5),color='red')
```

5 Model de Verhulst (II)

Els primers punts fan que la recta trobada s'allunyi dels altres punts. Usant més informació referent a la situació que es vol modelar es poden plantejar algunes modificacions, tant pel que fa a les dades com al model. Una situació excepcional com va ser la segona guerra mundial va pertorbar molt les dades reals per tant és raonable eliminar els primers tres valors de la taula.

Quina és la recta de regressió amb aquesta nova consideració?

Quin és el coeficient de correlació?

Quin és l'error comès?

Si $t_0 = 1965$ i $p_0 = 3.342771$ quina és la expressió de $p(t)$? Tot i que la equació diferencial $p' = p(ap + b)$ és fàcil de resoldre ja que és de variables separades, `sagemath` no se'n surt amb 'desolve.' L'hem de donar un cop de mà:

```
var('a,b,y'); integral(1/(y*(a*y+b)),y);
sols=solve([log(y/(a*y+b))-log(Yt[3]/(a*Yt[3]+b))==b*(t-Xt[3])],y);sols
```

Omple la taula de valors calculats per $p(t)$, usant l'expressió anterior:

t	1965	1970	1975	1980	1985	1990	1995	2000	2005
$p_{obs}(t)$	3.3428	3.6987	4.0761	4.4518	4.8553	5.2949	5.7191	6.1241	6.5148
$p_{mod}(t)$									

Fes la gràfica dels valors calculats i els observats per veure que el model s'ajusta força bé.

Extrapola la població mundial en els propers anys:

t	2010	2015	2020	2025	2030	2035	2040	2045	2050
$p_{mod}(t)$									

També es poden fer prediccions inverses, per exemple: quan vam arribar als 7 mil milions.

Fes la gràfica per un interval més gran per veure la tendència del model.

A partir de la gràfica es veu que la població mundial tendeix a estabilitzar-se. A quin valor? Què val $\lim_{t\to\infty} p(t)$?

6 2005 versus 2020

Cada any es publica[2] una revisió de les dades presentades a l'inici. Compara els resultats obtinguts amb els diversos models amb les darreres dades publicades per veure com el teu model s'ajusta a la realitat. En particular compara la predicció amb les dades actuals donada pels tres models estudiats amb les dades fins al 2005 amb la donada per oficial durant la revisió actual. Com variarien les prediccions pels propers 50 anys si féssim servir les noves dades? T'atreviries a predir quan arribarem als 8 mil milions?

De fet, és d'aquesta manera com els científics poden predir fins i tot el dia de l'any (i quasi be fins i tot el minut) en que neix aquest habitant de dígit tan especial. Un altre cosa és encertar el país on passa. Si més no, un cop es sap el minut en que passarà, molts països tracten d'atribuir-se el "mèrit" adjudicant-li el títol a algú nascut a prop d'aquell minut, que sempre n'hi ha algun.

Fes una taula comparativa dels valors que s'obtenen a l'extrapolar usant els tres models, tant amb les dades originals com les modificades.

t	$p_{mod1}(t)$	$p_{mod2}(t)$	$p_{mod3}(t)$
2010			
2015			
2020			
2025			
2030			

t	$p_{mod1}(t)$	$p_{mod2}(t)$	$p_{mod3}(t)$
2035			
2040			
2045			
2050			
2055			

Per tal de que cadascú de vosaltres faci quelcom diferent, agafeu lin=mod(NIU,263)+42 i prediu la població a l'any 2050 del país (o territori) que ocupa la línia "lin" en l'arxiu Excel que apareix en primer lloc del enllaç de la UN (Download Data Files). Preneu com a dades inicials els mateixos anys 1950-2020 i decidiu en funció del país, quin rang d'anys és apropiat agafar pel model de Verhulst.

Determineu també el nombre d'habitants d'aquell país o territori que hi habia el dia del vostre neixement.

Si us toca un pais o territori amb una població molt baixa, o hi detecteu un comportament caòtic, podeu escollir un altre territori proper a la taula.

[2] `http://esa.un.org/unpd/wpp/Excel-Data/population.htm`

7 Estadística i correlació

Recordeu d'estadística que, donat un conjunt de dades $X = [X_i, i = 0, \ldots, n], Y = [Y_i, i = 0, \ldots, n]$, la mitjana μ_X de X, la desviació estàndard σ_X de X i la correlació $cor(X,Y)$ són definides per

$$\mu_X = \frac{1}{n+1}\sum_{i=0}^{n} X_i$$

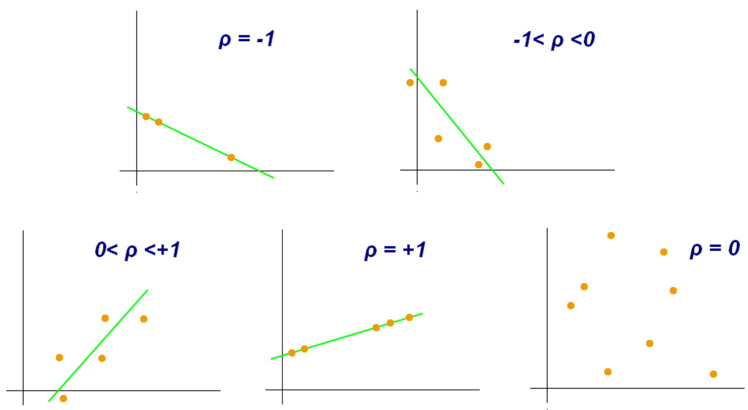
$$\sigma_X^2 = \frac{1}{n}\sum_{i=0}^{n}(X_i - \mu_X)^2$$

$$cor(X,Y) = \frac{1}{n}\sum_{i=0}^{n} \frac{X_i - \mu_X}{\sigma_X} \cdot \frac{Y_i - \mu_Y}{\sigma_Y}$$

En `sagemath` tenim programades la mitjana muX i la desviació estàndard sigmaX, però no la correlació:

```
muX=mean([X[i] for i in range(0,len(X))]);
sigmaX=std([X[i] for i in range(0,len(X))]);
```

El coeficient de la correlació $\rho = \mathrm{cor}(X,Y)$ és un índex que mesura la relació lineal entre dues variables quantitatives. El valor de l'índex de correlació varia en l'interval $[-1,+1]$. Si $\rho = 1$, hi ha una correlació positiva perfecta, i l'índex indica una dependència total entre les dues variables anomenada relació directa: quan una d'elles augmenta, l'altra també ho fa en proporció constant. Si $0 < \rho < 1$, hi ha una correlació positiva. Si $\rho = 0$, no existeix relació lineal. Però això no necessàriament implica que les variables són independents: poden existir encara relacions no lineals entre les dues variables. Si $-1 < \rho < 0$, hi ha una correlació negativa. Si $\rho = -1$, hi ha una correlació negativa perfecta, i l'índex indica una dependència total entre les dues variables anomenada relació inversa: quan una d'elles augmenta, l'altra disminueix en proporció constant.



8 Informe

Per ser la primera pràctica, he fet a classe la major part de la tasca que estava prevista a fer com a exercici a entregar.

De fet, possiblement només ha quedat per fer l'apartat 6 amb les dades actualitzades.

Per tant, la tasca que us encarrego és la de donar els resultats que es pregunten a l'apartat 6.

Si disposéssim d'un calendari més espaiat de pràctiques, podria donar un parell de setmanes per entregar aquesta pràctica. Però al tenir cinc pràctiques seguides haig de recomanar molt fortament que es facin setmanalment.

Feu l'entrega a través del Campus Virtual. Entregueu un fitxer SAGE amb les instruccions, resultats i comentaris que volgueu fer.