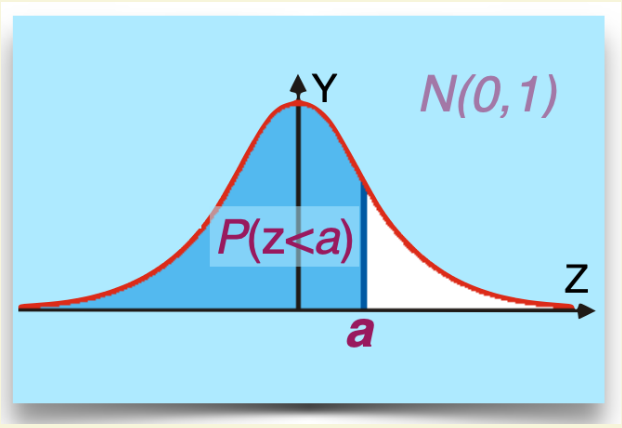
Lliurament 8:
Distribucions de probabilitat

Matemàtiques II

Josep Mulet Pol

Àmbit científic

IEDIB



IEDIB

https://iedib.net/

Edició LaTeX: ® **Josep Mulet Pol**

Versió: 27-01-2023

Índex

1. Concepte de variable aleatòria

Concepte de variable aleatòria

Considerem un experiment aleatori, per exemple, S=llançar una moneda a l'aire. L'espai mostral està format per tots els possibles esdeveniments elementals i val $\Omega = \{C, X\}$

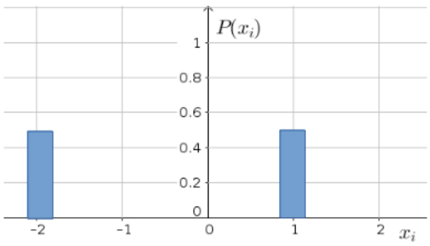
> Es defineix **variable aleatòria** a la llei (o funció) que assigna a cada resultat d'un experiment aleatori **un nombre real** .

Podríem definir una variable aleatòria de la següent forma; si surt cara guanyes un euro; si surt creu perds 2 €

$$\begin{array}{cccc} x: & \Omega & \to & \mathbb{Z} \\ & C & \to & 1 \\ & X & \to & -2 \end{array} \tag{1}$$

En aquest joc (que no és just perquè acabes perdent diners), els valors que pren la variable x són $x_i = \{-2, 1\}$, tots dos amb igual probabilitat. Assignar probabilitats a cada valor de la variable aleatòria es coneix com construir la **distribució de probabilitat** . Aquesta es pot expressar en forma de taula o com una gràfica:

x_i	$P(x_i)$
-2	$\frac{1}{2}$
1	$\frac{1}{2}$



En l'exemple anterior, els valors de la variable aleatòria són discrets, i per tant, diem que es tracta d'una **distribució de probabilitat discreta** .

> Si $x = \{x_1, \cdots, x_m\}$ és una variable aleatòria discreta, la seva distribució de probabilitat discreta s'ha de complir:
>
> - $P(x_i) \geq 0$. Les probabilitats són positives o zero.
> - $\sum P(x_i) = 1$. La suma de totes les probabilitats han d'ésser igual a 1.

Considerem un altre experiment S=”elegir a l'atzar una persona d'una classe i mesurar-ne l'altura”. La llei que associa a cada alumne la seva altura és una variable aleatòria. Ara bé, la variable $x = altura$ pot prendre qualsevol valor dins d'un interval de la recta real i, per tant, l'anomenam **distribució de probabilitat contínua** .

En aquest lliurament estudiarem un tipus de distribució de probabilitat discreta (la binomial) i una de contínua (la distribució normal).

Vídeo 8.1: *Concepte de variable aleatòria*
https://www.youtube.com/watch?v=CHuJtQ0CRfs

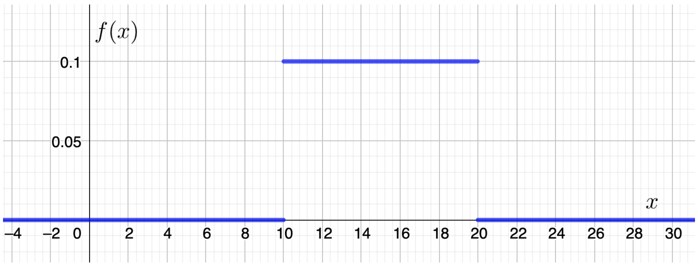
2. Distribucions de variable contínua

Considera l'experiment aleatori anar de casa al treball. Per a cada dia que realitzes aquest experiment mesures el temps que tardes a arribar-hi. El temps és una **variable contínua** perquè pot valer qualsevol valor dins d'un interval. Vegem un exemple molt senzill de com es descriu aquesta variable aleatòria.

Emprarem una distribució uniforme per resoldre aquest problema. Això vol dir que tots els temps entre 10 i 20 minuts tindran igual probabilitat i, tots els altres temps tenen probabilitats zero. La funció de probabilitat ve donada per la següent funció:

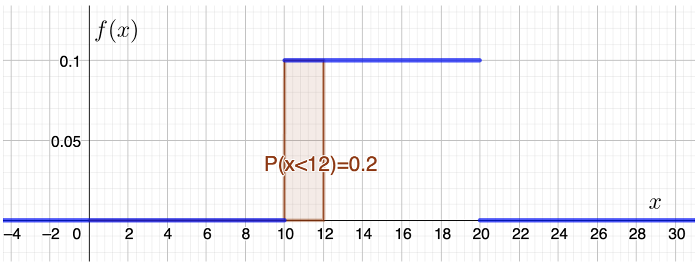
$$f(x) = \begin{cases} 0.1 & \text{si } 10 \leq x \leq 20 \\ 0 & \text{en altre cas} \end{cases} \tag{2}$$

La seva gràfica es mostra a continuació i es fàcil comprovar que sempre és positiva o zero i que **l'àrea davall** d'ella val 1. Compleix les propietats d'una funció de distribució de probabilitat.



Com es calculen les probabilitats?

Si volem saber la probabilitat d'arribar al treball amb menys de 12 minuts, el que hem de fer és calcular l'àrea davall la funció en l'interval $(-\infty, 12)$. Aquesta àrea s'obté de l'àrea del rectangle colorejat en la següent figura i val $P(x < 12) = 0.1 \cdot (12 - 10) = 0.2$.



En canvi, si volem saber la probabilitat d'arribar al treball exactament en 12 minuts, la probabilitat és zero, $P(x = 12) = 0$, perquè no es forma cap àrea davall la funció.

Si volem saber la probabilitat que arribem al treball en un temps exactament igual a un valor $P(x = a)$, la resposta és zero per a qualsevol variable aleatòria contínua.

Exercicis

1. S'ha comprovat que el temps d'engegada d'un portàtil segueix una distribució uniforme entre 8 i 16 segons. Escriu l'equació de la funció de

distribució de probabilitat i calcula la probabilitat que el portàtil tardi més de 10 segons a engegar-se.

2.1 La distribució normal

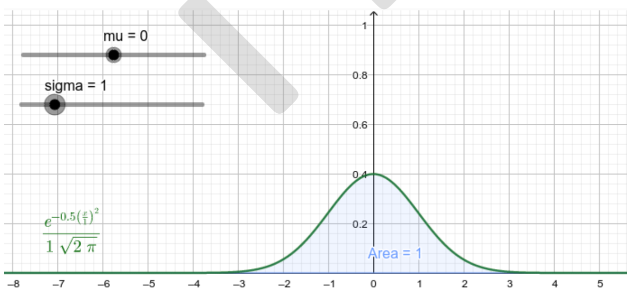
La **distribució normal** és una distribució de probabilitat **contínua** i s'anomena així perquè apareix en moltíssims fenòmens relacions amb psicologia, pedagogia, biologia... Tot i que no tots els fenòmens es descriuen per una normal, en bon grau s'hi aproximen a ella.

Llavors, "elegir a l'atzar una persona d'una classe i mesurar-ne l'altura" vindrà descrit per una distribució que, en bona d'aproximació, considerarem una distribució normal.

Matemàticament, aquesta distribució es descriu mitjançant una funció $N(x)$ que té forma de çampana". Algunes propietats d'aquesta funció són:

- $N(x)$ sempre és positiva.
- $N(x)$ és simètrica respecte del seu màxim.
- L'àrea sota la corba $N(x)$ és igual a 1.

En realitat, no existeix una única funció normal $N(x)$ sinó una família d'elles segons la posició del màxim i la seva amplada. Definim μ (mitjana) com la posició del màxim i σ (desviació típica) al paràmetre que controlarà l'amplada. En la següent simulació podeu modificar aquests dos paràmetres i observar com canvia la corba. Recordeu, però, que l'àrea davall la corba sempre és igual a 1.



Simulació 2: https://www.geogebra.org/m/cwq3hbkd : *Modifica la mitjana i la desviació típica i observa com canvia la corba normal.*

Vídeo 8.2: *La distribució normal*
https://www.youtube.com/watch?v=cr31rFZ87HA

Càlcul de probabilitats

Si ens demanen quina és la probabilitat que en triar un alumne la seva altura sigui exactament $x = 170$ cm, pot semblar una mica estrany però, la resposta és zero. Si volem una altura exacta de $x = 170$ cm, hem d'aconseguir que totes les xifres decimals de la mesura tinguin un valor determinat, cosa que és pràcticament impossible. D'aquí trobam la primera regla:

> En una distribució contínua qualsevol, les probabilitats d'un valor en concret són zero, $P(x = a) = 0$

En realitat, estarem interessats a conèixer probabilitats d'intervals. Per exemple, la probabilitat que l'altura estigui entre compresa entre 165 i 175 cm. Les probabilitats d'un interval es troben amb la segona regla:

> En una distribució contínua, les probabilitat en un interval $[a, b]$ s'obté de l'àrea davall la funció $N(x)$ en aquell interval.

2.2 Càlcul de probabilitats amb la distribució normal

A l'apartat anterior hem comentat les dues regles per calcular probabilitats:

> - $P(x = a) = 0$
> - $P(a \le x \le b) = \int_a^b N(x)\,dx$

L'única dificultat en aquest procès, però, és que és impossible calcular la integral de la corba normal $N(x)$ de forma exacte (no es coneix el resultat de la primitiva). Això ens obliga a seguir un altre camí; primer aprendrem a calcular probabilitats a partir d'una taula de valors i, tot seguit, generalitzarem el procés a una corba normal qualsevol.

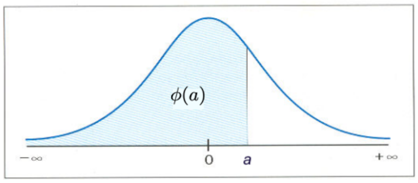
La normal estàndard

La **normal estàndard** és aquella corba normal que té mitjana zero ($\mu = 0$) i desviació típica 1 ($\sigma = 1$). Per aquesta normal, existeix una taula que ens proporciona la probabilitat que la variable (que ara anomenarem z) sigui menor que un valor a: $P(z \le a)$. A partir d'ara escriurem indistintament

$P(z \leq a)$ o $P(z < a)$ perquè, com ja s'ha dit, la probabilitat d'un valor en concret és zero.

Per deixar clara la diferència, a les probabilitats que obtinguem a partir de la taula, les anomenarem $\phi(a) = P(z \leq a)$

$\phi(a) = P(z \leq a)$



a	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Vegem exemples com s'utilitza aquesta taula

Exemple 1

Cerca a la taula $P(z \leq 1.6)$ i $P(z \leq 1.97)$

El valor de les unitats i dècimes es cerca a la columna vertical. El valor de les centèsimes a la fila horitzontal. L'element de la matriu és el valor de la probabilitat.

a	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	$\boxed{0,9452}$	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	$\boxed{0,9756}$	0,9761	0,9767

$P(z \leq 1.6) = 0.9452$ i $P(z \leq 1.97) = 0.9756.$

Pot passar que estiguem interessats amb un valor que no surt a la taula, com ara $P(z \leq 1.975)$. En tal cas s'agafa la mitjana entre $P(z \leq 1.97)$ i $P(z \leq 1.98)$ que sí es troben a la taula.

Per trobar la probabilitat en un interval utilitzarem el fet que les àrees són additives

$$P(a \leq z \leq b) = \phi(b) - \phi(a) \tag{3}$$

Per trobar la probabilitat que $z \geq a$, utilitzam el esdeveniment contrari

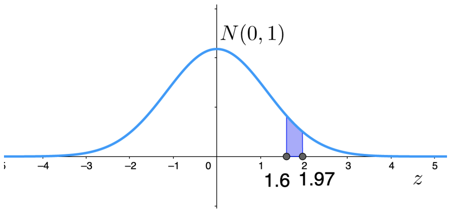
$$P(z \geq a) = 1 - \phi(a) \tag{4}$$

Exemple 2

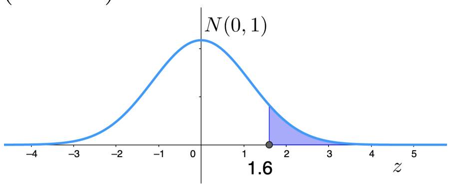
a) Cerca la probabilitat $P(1.6 \leq z \leq 1.97)$
b) Cerca la probabilitat $P(z > 1.6)$

a) A l'exemple anterior havíem trobat $P(z \leq 1.6) = \phi(1.6) = 0.9452$ i $P(z \leq 1.97) = \phi(1.97) = 0.9756$.

$$P(1.6 \leq z \leq 1.97) = \phi(1.97) - \phi(1.6) = 0.9756 - 0.9452 = 0.0304$$



b) $P(z > 1.6) = 1 - P(z < 1.6) = 1 - 0.9452 = 0.0584$



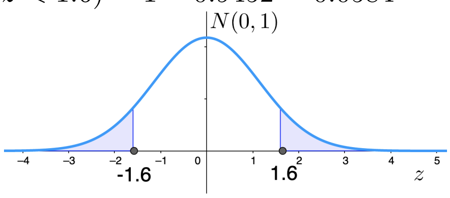
Finalment, si ens donen un valor de z negatiu, ens adonam que no apareix a la taula. Aquí utilitzam el fet que la corba $N(x)$ és simètrica respecte $x = 0$

$$P(z < -a) = P(z > a) = 1 - P(z < a) \tag{5}$$

Exemple 3

Cerca la probabilitat $P(z < -1.6)$

$P(z < -1.6) = 1 - P(z < 1.6) = 1 - 0.9452 = 0.0584$



Vídeo 8.3: *La distribució normal estàndard*
https://www.youtube.com/watch?v=92YAmsciM74

Exercicis

2. Suposant que la variable aleatòria z segueix una distribució normal estàndard, calcula les següents probabilitats:

a) $P(z < 1,5)$

b) $P(z > 1,25)$

c) $P(1,3 < z < 1,37)$

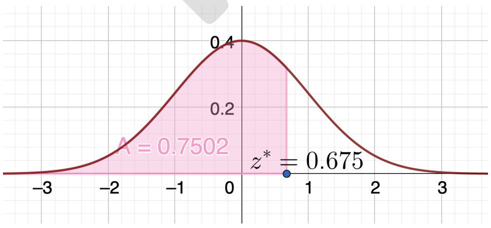
d) $P(z < -0,75)$

e) $P(z \geq -1,34)$

2.3 Cercar z sabent la probabilitat

En aquesta secció continuam treballant amb la taula de probabilitats de la normal estàndard $N(0;1)$ però en aquesta ocasió l'objectiu no serà trobar la probabilitat sinó la z per la qual s'obté una probabilitat. És a dir, és el problema invers de l'explicat a l'apartat anterior.

Suposau que necessitam saber quina z^* deixa per davall el 75% de la distribució. En altres paraules, quina z^* compleix que $P(z < z^*) = 0.75$. Gràficament, significa on hem de situar la z^* perquè l'àrea marcada valgui 0.75.



Ara, hem de cercar a la taula el valor de probabilitat que més s'acosti a 0.75 i, una vegada localitzat, mirar per quina z passa.

a	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389

La probabilitat 0.75 no apareix en la taula però hi ha un canvi d'un valor menor a un major que 0.75 entre $z^* = 0.67$ i $z^* = 0.68$. En primera aproximació, podem prendre el valor intermedi, $z^* = 0.675$.

Exemple 4

Quina z^* deixa per sobre 5% de la població en una distribució normal estàndard?

Ens demanem per a quina z^* es compleix que $P(z > z^*) = 0.05$ per a la distribució $N(0; 1)$.
Per comoditat, pensarem que per davall ha de quedar el 95% de la població, és a dir, $P(z < z^*) = 0.95$.
Ara consultam a la taula de normal una probabilitat de 0.95 i veim que s'aconsegueix per a un bastant proper a $z^* = 1.65$.

a	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767

Exercicis

3. Calcula el valor de z^* en la distribució $N(0, 1)$ per al qual es compleix:

a) $P(z < z^*) = 0.95$

b) $P(z > z^*) = 0.01$

c) $P(-z^* < z < z^*) = 0.90$

2.4 Tipificació de la variable

En l'apartat anterior hem après a calcular probabilitats amb la normal estàndard. A continuació, aprendrem a calcular probabilitats amb altres corbes normals. Considerem, per tant, una normal $N_{\mu,\sigma}(x)$ de mitjana μ i desviació típica σ.

Tipificació de la variable

Passarem de la variable x descrita per la distribució $N_{\mu,\sigma}(x)$ a la variable z

$$z = \frac{x-\mu}{\sigma} \tag{6}$$

que segueix la normal estàndard $N_{1,0}(z)$ de la qual ja sabem calcular probabilitats.

Aquest exemple il·lustra els passos a seguir:

Exemple 5

L'altura d'un grup d'estudiants x es modela mitjançant una normal de mitjana 165 cm i desviació 10 cm. Calcula la probabilitats:
a) $P(x < 172)$
b) $P(160 < x < 172)$

Passam a tipificar la variable x

$$x = 160 \quad \rightarrow \quad z = \frac{160-165}{10} = -0.5 \tag{7}$$

$$x = 172 \quad \rightarrow \quad z = \frac{172-165}{10} = 0.7 \tag{8}$$

Les probabilitats que ens demanen es poden expressar com

a) $P(x < 175) = P(z < 0.7) = \phi(0.7) = 0.7580$

b) $P(160 < x < 172) = P(-0.5 < z < 0.7) = P(z < 0.7) - P(z < -0.5) =$
$P(z < 0.7) - [1 - P(z < 0.5)] = \phi(0.7) + \phi(0.5) - 1 = 0.7580 + 0.6915 - 1 = 0.4495$

Vídeo 8.4: *Tipificació de la variable*
https://www.youtube.com/watch?v=Oc05EaiHp8k

Càlcul de la variable sabent la probabilitat

Seguim amb l'exemple anterior. Suposau que ara necessitam dividir els alumnes en dos grups de forma que en el primer grup hi hagui el 70% dels alumnes més baixos i a l'altre grup la resta d'alumnes. Per fer-ho, necessitam saber a partir de quina altura x^* direm que l'alumne vagi al segon grup.

Ens estam demanant per a quina x^* queda per davall el 70% de la població $P(x < x^*) = 0,7$

Si tipificam la variable, $z^* = \dfrac{x^* - 165}{10}$, volem cercar el valor de z^* en la taula de la normal estàndard que compleix $P(z < z^*) = 0,7$.

Vegem com fer-ho damunt la taula. Primer cercam la probabilitat que més s'apropi a 0,7 i després quina z li correspon. Si no trobam el valor de probabilitat $0,7$, farem una mitjana dels valors z més propers si és necessari.

a	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389

Veim que això passa per a z^* entre 0,52 i 0,53; per tant, pendrem $z^* = 0,525$. Substituint el valor de z^* trobat dins $z^* = \dfrac{x^* - 165}{10}$ i resolent l'equació

$$0,525 = \frac{x^* - 165}{10} \quad \rightarrow \quad \rightarrow x^* = 165 + 10 \cdot 0,525 = 170,25 \text{ cm} \tag{9}$$

Ara ja sabem que els alumnes més baixos que $170,25$ cm han d'anar al primer grup i la resta al segon.

Exercicis

4. Per a la normal de mitjana 173 i desviació típica 6, calcula les probabilitats

a) $P(x \leq 173)$

b) $P(x \leq 170)$

c) $P(x \geq 179)$

d) $P(176 \leq x \leq 179)$

5. El pes d'una població segueix una distribució normal de mitjana 70 kg i desviació típica desconeguda. La probabilitat que a l'agafar un individu a l'atzar pesi més de 85 kg és del 10%. Quina és la desviació estàndard de la distribució?

2.5 Problemes PBAU distribució normal

Exemple 6

JUNY 2017

El test d'intel·ligència (CI) és una prova que en teoria mesura la intel·ligència de l'individu i dona un valor que aproximadament té de mitjana 100. És a dir, el nivell 100 se suposa que és el nivell d'intel·ligència d'una persona normal. Suposem ara que el nivell d'intel·ligència d'una determinada població segueix una distribució normal de mitjana 100 i desviació típica 10.

a) Calculau el percentatge de la població que es considera superdotada. Una persona es considera superdotada si té un nivell d'intel·ligència superior a 130.

b) Calculau el percentatge de la població amb un nivell d'intel·ligència entre 90 i 110.

c) Ens diuen que el 70% de la població té un nivell d'intel·ligència menor que un cert llindar. Calculau aquest llindar.

Sigui X la variable aleatòria que ens dona el nivell d'intel·ligència d'un individu qualsevol de la població.

a) Ens demanen $P(X \geq 130)$. Si tipificam la variable, obtenim:

$$z = \frac{130 - 100}{10} = 3 \tag{10}$$

$$P(X \geq 130) = P(z \geq 3) = 1 - P(z < 3) = 1 - 0.9987 = 0.0013 \tag{11}$$

on Z és una normal estàndard.
Per tant, el 0.13% de la població és superdotada.

b) Ens demanen $P(90 \leq X \leq 110)$; tipificam la varible:

$$z = \frac{90 - 100}{10} = -1 \quad z = \frac{110 - 100}{10} = 1 \tag{12}$$

$$P(90 < X < 110) = P(-1 \leq z \leq 1) = P(Z \leq 1) - P(Z \leq -1) =$$

$$P(Z \leq 1) - [1 - P(Z \leq 1)] = 0.8413 - 0.1587 = 0.6827 \tag{13}$$

El percentatge serà: 68.27%.

c) Ens diuen que $P(Z \leq z) = 0.7$, on z és el llindar a calcular. Si cercam a la taula el valor de z pel qual la probabilitat és 0.7, es troba entre 0.52 i 0.53, agafam la mitjana 0.525. Tot seguit necessitam trobar la x. Escrivim la fórmula de la tipificació:
Tenim l'equació $\frac{x - 100}{10} = 0.525$, de la qual és senzill aïllar $x = 105.25$.

Exemple 7

SETEMBRE 2017

El temps que un alumne pot estar concentrat i escoltar el professor en una classe de Matemàtiques es modela com una distribució normal de mitjana 15 minuts i desviació típica 5 minuts.

a) Calculau la probabilitat que un alumne estigui concentrat més de 20 minuts.

b) Calculau la probabilitat que un alumne estigui concentrat entre 10 i 30 minuts.

c) Ens diuen que la probabilitat que un alumne estigui concentrat més de x minuts val 0.75. Calculau aquest valor de x minuts.

Sigui X la variable aleatòria que ens dona el temps que està concentrat un alumne en una classe de Matemàtiques.

a) Ens demanen $P(X > 20)$. Si tipificam la variable, obtenim:

$$z = \frac{20 - 15}{5} = 1 \tag{14}$$

Llavors, expressam la probabilitat com

$$P(X > 20) = P(Z > 1) = 1 - P(Z < 1) = 1 - 0.8413 = 0.1587 \tag{15}$$

on Z és una normal estàndard

b) Ens demanen $P(10 < X < 30)$:

$$z = \frac{10 - 15}{5} = -1 \tag{16}$$

$$z = \frac{30 - 15}{5} = 3 \tag{17}$$

$$P(10 < X < 30) = P(-1 < Z < 3) = P(Z < 3) - P(Z < -1) =$$

$$P(Z < 3) - [1 - P(Z < 1)] = 0.9987 - 0.1587 = 0.84 \tag{18}$$

c) Ens diuen que $P(X > x) = 0.75$, o si es vol $P(X < x) = 1-0.75 = 0.25$. Cercam a la taula de la normal estàndard per a quin valor de z la probabilitat assoleix 0.25. Trobam un valor aproximat de $z = -0.675$. Tipificam:

$$-0.675 = \frac{x - 15}{5} \tag{19}$$

i aïllam el valor de $x = 15 - 5 \cdot 0.675 = 11.627$.

Exemple 8

JUNY 2018

Considerem la població d'estudiants que han aprovat la selectivitat en la convocatòria de juny un any determinat. Sigui X la variable aleatòria que modela la proporció d'estudiants (tant per u) de la població anterior que escull estudiar un grau d'humanitats. Aquesta variable aleatòria X es modela amb una distribució normal de mitjana 0.35 i desviació típica 0.1. Es demana:
a) Quina és la probabilitat que en un any qualsevol més del 45% dels estudiants de la població considerada estudiïn un grau d'humanitats? b) En els darrers 10 anys, en quants anys el percentatge d'estudiants de la població considerada que han escollit estudiar un grau d'humanitats no ha superat el 30%?

a) Demana calcular $P(X > 0.45)$. Tipificam

$$z = \frac{0.45 - 0.35}{0.1} = 1 \tag{20}$$

Expressam la probabilitat com

$$P(X > 0.45) = P(Z > 1) = 1 - P(Z < 1) = 1 - 0.8413 = 0.1587 \tag{21}$$

b) Ens demanen $P(X < 0.3)$. Tipificam

$$z = \frac{0.3 - 0.35}{0.1} = -0.5 \tag{22}$$

Expressam la probabilitat com

$$P(X < 0.3) = P(Z < -0.5) = 1 - P(Z < 0.5) = 1 - 0.6915 = 0.3085 \tag{23}$$

o el 30.85% dels casos. En deu anys, tindrem 30.85% de $10 = 3.085$ anys. La resposta és en 3 anys es dóna la condició que la matrícula no ha superat el 30%.

Exemple 9

JULIOL 2018
El nombre de passes que fa el professor *Jaimito* durant una hora de classe es modela amb una distribució normal de mitjana 100 passes i desviació típica 20.5 passes.
a) Calculau la probabilitat que el professor faci més de 125 passes durant una classe.
b) Ens diuen que en el 45% de les classes que fa el professor aquest fa menys de x passes. Trobau aquest valor x.

Sigui X la variable aleatòria que ens dóna el nombre de passes que fa el professor durant una hora de classe. Ens diuen que $X : N(\mu = 100, \sigma = 20.5)$. Ens demanen:

a) $P(X > 125) = P(Z > 1.22) = 0.1112$, on Z representa la distribució normal estàndard.

b) Sabem $P(X < x) = 0.45$. Cercam a la taula de la normal estàndard per a quin valor de z la probabilitat assoleix 0.45. Donat que no el trobam, cercarem $1 - 0.45 = 0.55$ i sabem que z serà negativa. Trobam un valor aproximat de $z = -0.125$. Tipificam:

$$-0.125 = \frac{x - 100}{20.5} \tag{24}$$

i aïllam el valor de $x = 100 - 20.5 \cdot 0.125 = 97.44$ passes.

Exercicis

6. La despesa total diària d'una família és una variable normal amb mitjana 36€ i desviació típica 4€.

a) Quina és la probabilitat que en un dia es gastin més de 36€?

b) Quin percentatge de dies la despesa està entre 32€ i 36€?

c) Calcula la despesa per davall de la qual es troba el 80% dels casos.

7. En una fàbrica de cotxes es fan proves per a conèixer el temps que tarden els seus vehicles a aconseguir la velocitat punta. Es considera aquest temps com una variable normal de mitjana 20 s i desviació típica 2 s. Calcula:

a) La probabilitat que un vehicle assoleixi la seva velocitat punta en 25 s.

b) La probabilitat que assoleixi la velocitat punta amb menys de 18 s.

c) La probabilitat que assoleixi la velocitat punta entre 18 s i 22 s.

d) Volem classificar els cotxes en lents i ràpids. Volem que només el 10% dels cotxes siguin considerats ràpids. A partir de quina velocitat podem dir que un cotxe és ràpid?

3. Distribucions de variable discreta

Considerem el joc de la moneda; si surt cara guanyes un euro; si surt creu perds 2 €

$$\begin{array}{rccc} x: & \Omega & \rightarrow & \mathbb{Z} \\ & C & \rightarrow & 1 \\ & X & \rightarrow & -2 \end{array} \tag{25}$$

Sabem que la variable aleatòria diners guanyats/perduts en una partida poden prendre dos valor $x_i = \{-2, 1\}$ amb probabilitats $P(x = -2) = \frac{1}{2}$, $P(x = 1) = \frac{1}{2}$.

Esperança o mitjana

Ara ens demanam, si jugam moltes vegades a aquest joc, quants euros en promig guanyarem o perdrem? Aquest concepte es coneix com **esperança matemàtica** .

Mitjana o Esperança matemàtica

$$\mu(x) = \sum x_i P(x = x_i) = x_1 \cdot p_1 + \cdots + x_m \cdot p_m \tag{26}$$

Exemple 10

Per al joc de les monedes,

$$\mu(x) = \sum x_i P(x = x_i) = -2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = -0.5 \tag{27}$$

Aquesta esperança significa que si jugam el joc 1000 vegades, perdrem en mitjana $-0.5 \cdot 1000 = -500$ €. Com ben bé diu la dita, “La banca sempre guanya”.

Variància

De l'estadística, sabem que la mitjana ens dona el valor central de la distribució però no ens informa de la dispersió o variabilitat dels valors. Per això, empram dos paràmetres anomenats variància i desviació típica.

Variància:

$$Var(x) = \sum x_i^2 P(x = x_i) - \mu(x)^2 \tag{28}$$

Desviació típica:

$$\sigma(x) = \sqrt{Var(x)} \tag{29}$$

Exemple 11

Per al joc de les monedes, la variància és

$$Var(x) = (-2)^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} - (-0.5)^2 = 2.25 \tag{30}$$

La desviació típica és

$$\sigma(x) = \sqrt{2.25} = 1.5 \tag{31}$$

El significat de la desviació típica és el següent. Si consideram l'interval $[\mu - \sigma, \mu + \sigma] = [-0.5 - 1.5, -0.5 + 1.5] = [-2, 1]$, podem afirmar que el guany que esperem que s'obtingui en una partida és entre -2 i 1 €. Com veim, ens informa de la variabilitat dels valors que pot pendre la variable aleatòria.

3.1 Experiència dicotòmiques

La distribució binomial és una distribució **discreta** , és a dir, la variable aleatòria pren valors enters. La distribució binomial s'aplica a **experiència dicotòmiques** repetida n vegades. Vegem que significa això.

Experiència dicotòmica

Si d'una experiència només ens fixam si ocorre un esdeveniment A o no ocorre ($\overline{A}$); es tracta d'una experiència dicotòmica (només passen dues coses).

Direm **èxit** que passi el esdeveniment A i **fracàs** que no passi. Anomenarem p la probabilitat d'èxit i $q = 1 - p$ la probabilitat de fracàs.

Alguns exemples de variables dicotòmiques són

- Llançar una moneda i mirar si surt C: èxit $\to$ cara, $p = 0.5$; fracàs $\to$ creu, $q = 0.5$.

- Llançar un dau i mirar si surt un sis: èxit $\to$ sortir un sis, $p = \frac{1}{6}$; fracàs $\to$ diferent a sis, $q = 1 - \frac{1}{6} = \frac{5}{6}$.

- Treure una carta i mirar si és un rei: èxit $\to$ sortir rei, $p = \frac{4}{40}$; fracàs $\to$ diferent a rei, $q = 1 - \frac{4}{40} = \frac{36}{40}$.

Distribució binomial

Definim la variable aleatòria discreta x que pren els valors $\{0, 1, 2, \cdots, n\}$ com el nombre d'èxits que passen en repetir n vegades una experiència dicotòmica.

En general, una variable aleatòria discreta segueix una distribució binomial si

1. està associada a una experiència dicotòmica i x és el nombre d'èxits en les n repeticions
2. l'experiència dicotòmica es repeteix n vegades
3. les repeticions són **independents**

Alguns exemples de variables que segueixen una distribució binomial es donen a continuació

- Llançar una moneda 5 vegades i comptar quantes cares surten: poden sortir $x = \{0, 1, 2, 3, 4, 5\}$ cares. èxit=Treure cara, $N = 5$ repeticions, $p = 0.5$, $q = 0.5$.

- Llançar un dau 2 vegades i mirar quants sis surten: poden sortir $x = \{0, 1, 2\}$ sis. èxit=Treure sis, $N = 2$ repeticions, $p = \frac{1}{6}$, $q = \frac{5}{6}$.

- Treure tres carta i mirar el número de reis: poden sortir $x = \{0, 1, 2, 3\}$ reis. èxit=Treure rei, $N = 3$ repeticions, $p = \frac{1}{10}$, $q = \frac{9}{10}$.

Vídeo 8.5: *La distribució binomial*
https://www.youtube.com/watch?v=1gWD79mvPK4

3.2 Repàs de combinatòria

En aquesta secció aprendrem a calcular probabilitats amb la distribució binomial, però, primer necessitam introduir una sèrie de conceptes.

Factorial d'un nombre

El factorial d'un nombre n positiu o zero l'escrivim com $n!$ és el producte d'ell amb tots els nombres menors a ell. Per definició $0! = 1$ i $1! = 1$. Alguns exemples són:

- $2! = 2 \cdot 1 = 2$
- $3! = 3 \cdot 2 \cdot 1 = 6$
- $7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$
- $\ldots$

Nombres combinatoris

El nombre combinatori $C(n, k)$ o també escrit com $\begin{pmatrix} n \\ k \end{pmatrix}$ és el nombre de maneres que es poden tenir k èxits en repetir un experiment n vegades.

Per exemple, de quantes maneres podem treure 2 cares en llançar una moneda 3 vegades?

Escrivim els casos: CCX, CXC, XCC. Hi ha 3 maneres.

De quantes maneres podem treure 2 sis en llançar un dau 4 vegades?

Escrivim els casos: SSXX, SXSX, SXXS, XXSS, XSXS, XSSX. Hi ha 6 maneres.

En general, els nombres combinatoris ens donen la resposta més ràpidament.

Aquesta és la definició:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \tag{32}$$

Comprovem que obtenim els mateixos resultats a partir d'aquesta fórmula

$\binom{3}{2} = \frac{3!}{2! \cdot 1!} = \frac{6}{2} = 3$

$\binom{4}{2} = \frac{4!}{2! \cdot 2!} = \frac{24}{2 \cdot 2} = 6$

Com veus, és més ràpid aplicar la fórmula que no pas fer el recompte de tots els casos. De quantes maneres podem treure 4 reis extraient 10 cartes d'una baralla amb reemplaçament? Fàcil

$\binom{10}{4} = \frac{10!}{4! \cdot 6!} = 210$ formes diferents.

Amb la calculadora pots emprar la tecla **nCr** per calcular directament el combinatori. En l'exemple anterior escrivim **10 nCr 4 =** i trobam **210** .

3.3 Càlcul de probabilitats amb la binomial

Si x és una variable que segueix una binomial $B(p; n$, amb probabilitat d'èxit p i n repeticions, la probabilitat d'obtenir k èxits és

$$P(x = k) = \binom{n}{k} p^k \cdot q^{n-k} \tag{33}$$

Exemple 12

Calculeu les següents probabilitats:
a) Probabilitat de treure 2 cares en llançar una moneda 3 vegades.
b) Probabilitat de treure 2 sis en llançar un dau 4 vegades.
c) Probabilitat de treure 4 reis en extreure 10 cartes amb reemplaçament.

a) Tenim que $n = 3$, $p = q = 0,5$, $P(x = 2) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} 0,5^2 \cdot 0,5^{3-2} = 0,375$

b) Tenim que $n = 4$, $p = 1/6$, $q = 5/6$, $P(x = 2) = \begin{pmatrix} 4 \\ 2 \end{pmatrix} (1/6)^2 \cdot (5/6)^{4-2} =$
$6 \cdot (1/6)^2 \cdot (5/6)^2 = 0,1157$

c) Tenim que $n = 10$, $p = 1/10$, $q = 9/10$, $P(x = 4) = \begin{pmatrix} 10 \\ 4 \end{pmatrix} 0,1^4 \cdot 0,9^{10-4} =$
$210 \cdot 0,1^4 \cdot 0,9^6 = 0,01116$

Exemple 13

Segons les estadístiques, el 18% de les dones espanyoles són fumadores. Si ens trobam a 7 dones pel carrer, calcula:
a) La probabilitat que cap d'elles sigui fumadora.
b) La probabilitat que una sigui fumadora.
c) La probabilitat que dues siguin fumadores.
d) La probabilitat que més de dues siguin fumadores.

Ens trobam davant una experiència dicotòmica F fumadora (èxit) i F' no fumadora (fracàs). La probabilitat d'èxit la dóna l'enunciat $p = 0,18$ i, per tant, $q = 1 - 0,18 = 0,82$. L'experiment es repeteix $n = 7$ vegades. LLavors, $x =$ n. de dones fumadores, segueix una distribució binomial.

a) $P(x = 0) = \begin{pmatrix} 7 \\ 0 \end{pmatrix} 0,18^0 \cdot 0,82^7 = 1 \cdot 0,82^7 = 0,2493$

b) $P(x = 1) = \begin{pmatrix} 7 \\ 1 \end{pmatrix} 0,18^1 \cdot 0,82^6 = 7 \cdot 0,18 \cdot 0,82^6 = 0,38305$

c) $P(x = 2) = \begin{pmatrix} 7 \\ 2 \end{pmatrix} 0,18^2 \cdot 0,82^5 = 21 \cdot 0,18^2 \cdot 0,82^5 = 0,25225$

d) Que hi hagi més de dos fumadors s'obté de $P(x > 2) = P(x = 3) + P(x = 4) + P(x = 5) + P(x = 6) + P(x = 7)$. Ara bé, serà més fàcil cercar el esdeveniment contrari $P(x > 2) = 1 - P(x \leq 2) = 1 - (P(x = 0) + P(x = 1) + P(x = 2))$ $= 1 - (0,2493 + 0,38305 + 0,25225) = 0,1154$

Es pot demostrar, no farem aquí, que la la mitjana i la desviació típica d'una variable binomial és

> Mitjana: $\mu = n \cdot p$
>
> Desviació típica: $\sigma = \sqrt{n \cdot p \cdot q}$

Vídeo 8.6: *Càlcul de probabilitats amb la distribució binomial*
https://www.youtube.com/watch?v=ku1ZYSzLVZk

Exercicis

8. En una distribució binomial amb paràmetres $p = 0,4$, $n = 10$, calcular: $P(x = 0)$, $P(x = 3)$ i $P(x = 9)$.

9. Tiram un dau 5 vegades. Calculau la probabilitat de treure 2 sis, 3 sis i 5 sis.

4. Aproximació normal de la binomial

PBAU

> Aquesta secció es deixa com material d'ampliació.
>
> Tot i que mai s'ha demanat a les PBAU, és un contingut de la programació de 2n de batxillerat.

Suposa que volem repetir l'exemple de les dones fumadores de la secció anterior, però ara considerant que hem demanat la pregunta a 200 dones, és a dir, $n = 200$. Recordem que la probabilitat d'èxit és $p = 0.18$ i fracàs $q = 1 - p = 0.82$. Ens demanam quina és la probabilitat que n'hi hagi exactament 45 fumadores?

Si aplicam la fórmula de la distribució binomial, caldrà calcular el següent

$$P(k = 45) = \binom{200}{45} 0.18^{45} \cdot 0.82^{155} \tag{34}$$

Si intentam fer el càlcul amb la calculadora ens trobam ja amb un problema, que el nombre combinatori $\binom{200}{45}$ és tan gran que la nostra calculadora no el pot donar. Obtenim un **Math Error!**

Per un moment, suposa que hagéssim pogut solucionar el problema anterior, però ara volem saber la probabilitat que hi hagi menys de 45 fumadors. Queda clar que per contestar aquesta pregunta caldria calcular 45 probabilitats amb la binomial i sumar-les totes; una feinada!

Per tots aquests motius s'empra el que es coneix com l'aproximació normal de la binominal. Quan el nombre de repeticions n és molt gran, la distribució binomial s'assembla molt a una normal i per això té sentit emprar la normal per calcular probabilitats. Vegem quan es pot fer:

Per poder aplicar l'aproximació normal de la binomial cal que

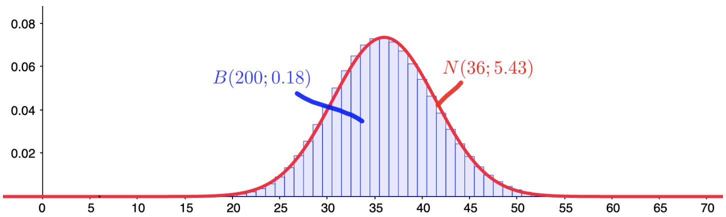
- $n > 30$
- $n \cdot p > 5$
- $n \cdot q > 5$

Per al problema anterior, es compleixen tots els requisits: $n = 200 > 30$, $n \cdot p = 200 \cdot 0.18 = 36 > 5$ i $n \cdot q = 200 \cdot 0.82 = 164 > 5$

Aleshores, la distribució binominal $B(200, 0.18)$ l'aproximam per la normal $N(\mu, \sigma)$ amb la mitjana i desviació de la binomial:

$$\mu = n \cdot p = 36 \qquad \sigma = \sqrt{n \cdot p \cdot q} = 5.43 \tag{35}$$

La gràfica següent mostra com les dues distribucions són molt semblants



Com que no podem calcular probabilitats de valors concrets amb la normal, caldrà cercar la probabilitat d'un interval centrat en el valor. Si volem saber $P(k = 45)$, haurem de calcular $P(44.5 < x < 45.5)$ amb la normal. Facem el càlcul:

Primer tipificam la variable

$$P(44.5 < x < 45.5) = P(x < 45.5) - P(x < 44.5) = P(z < 1.75) - P(z < 1.57)$$

Consultam els dos valors a la taula de la normal estàndard i obtenim la resposta:

$$= 0.9599 - 0.9418 = 0.0181$$

Exemple 14

Per a la situació descrita en aquesta secció, determinau la probabilitat que hi hagi menys de 45 dones fumadores.

Amb la binominal hauríem de calcular la probabilitat $P(k < 45) = P(k = 0) + P(k = 1) + \cdots + P(k = 44)$ cosa que és inviable.

Aplicarem l'aproximació normal amb la qual hem de calcular la probabilitat de què $P(x < 44.5)$. Recorda que el punt $k = a$, en la normal comença en $a - 0.5$ i acaba en $a + 0.5$.

$$P(x < 44.5) = P(z < \frac{44.5 - 36}{5.43}) = P(z < 1.57) = 0.9418$$

Vídeo 8.7: *L'aproximació normal de la binomial*
https://www.youtube.com/watch?v=acN-NcZsRYQ