

# Even herhalen...

We bespreken nu enkele specifieke veel voorkomende:

**KANSMODELLEN**

H2-H7

**POPULATIE**

H8

ONDERZOEKSVRAAG

H9-H13

STATISTIEK =  
op basis van een

**STEEKPROEF**

uit de populatie een  
uitspraak doen over de  
GANSE populatie

CONTEXT

# Statistiek: les 3

a. Binomiale verdeling

b. Normale verdeling (zie ook week 2)

 c. Lineaire combinaties (zie ook week 2) –  
het steekproefgemiddelde

d. Poisson verdeling

~~e. Exponentiële verdeling~~ (geen examenleerstof meer)

f. Benaderende verdelingen

Sabine Bertho

Giovanni Vanroelen

[sabine.bertho@kuleuven.be](mailto:sabine.bertho@kuleuven.be)

[giovanni.vanroelen@uhasselt.be](mailto:giovanni.vanroelen@uhasselt.be)

# Binomiaal kansexperiment: voorbeeld

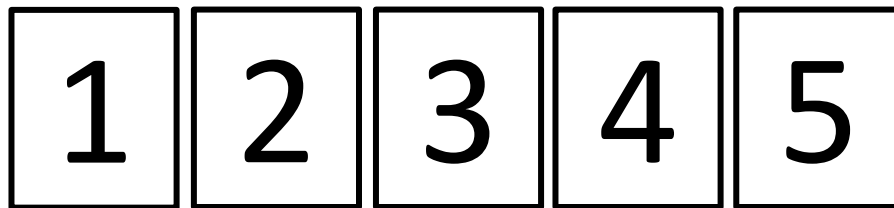
Je gooit 5x met een dobbelsteen. Wat is de kans dat je hierbij 2x een zes gooit?



- $P(\text{eerst 2x zes, daarna 3x geen zes}) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{7776}$

Maar eerst 3x geen zes en daarna 2x zes is ook ok! Of alle andere combinaties:

Plaatskaartjes:



$$C_5^2 = C_5^3$$

# Binomiaal kansexperiment: voorbeeld

Je gooit 5x met een dobbelsteen. Wat is de kans dat je hierbij 2x een zes gooit?



- $P(\text{eerst 2x zes, daarna 3x geen zes}) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{7776}$

Maar eerst 3x geen zes en daarna 2x zes is ook ok! Of alle andere combinaties:

$$P(2x \text{ een zes}) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot C_5^2 = \frac{125}{7776} \cdot nCr(5,2) = \frac{625}{3888} = 0,160751$$

2x succes      3x mislukking      mogelijke combinaties

# Binomiale kansexperiment: voorwaarden

Als een experiment aan de volgende voorwaarden voldoet, spreken we van een **binomiaal experiment**.

- Het experiment bestaat uit  $n$  deexperimenten die onderling onafhankelijk zijn.
- Elk deexperiment kan 2 waarden aannemen, die we noteren met 1 (*succes*) en 0 (*mislukking*).
- De kans op succes is voor elk deexperiment gelijk aan dezelfde kans  $p$ . De kans op mislukking, aangegeven met  $q = 1 - p$ , verandert dus ook niet.

Notatie:  $X \sim B(n, p)$

# Binomiaal kansexperiment: voorbeelden

Binomiaal experiment	Parameters	Binomiaal verdeelde variabele $X$
Gokken op 3 meerkeuzevragen, elk met 4 mogelijke antwoorden.	$n = 3, p = 1/4$	$X$ : <u>aantal</u> correcte antwoorden
Acht keer een muntstuk opgooien.	$n = 8, p = 1/2$	$X$ : <u>aantal</u> keer kop
Van een lading van 200 radio's weet men dat er 40 defect zijn. Je neemt er 10 willekeurig uit (met teruglegging) en controleert of ze al of niet defect zijn.	$n = 10, p = 0.2$	$X$ : <u>aantal</u> defecte radio's

# Binomiale kansfunctie: formularium en rekentoestel

Wanneer een kansexperiment bestaat uit  $n$  *onafhankelijke* deelexperimenten met elk **dezelfde** kans op succes ( $p$ ) spreken we van een **binomiaal kansexperiment**. De kansvariabele  $X$  = het aantal keer succes, volgt dan de binomiale kansverdeling:  $X \sim B(n, p)$ .

De bijhorende kansfunctie:  $P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$  ← rekentoestel: `binomPdf(n,p,i)`

Voor een binomiale verdeling geldt  $\mu_X = np$  en  $\sigma_X = \sqrt{np(1 - p)}$ .

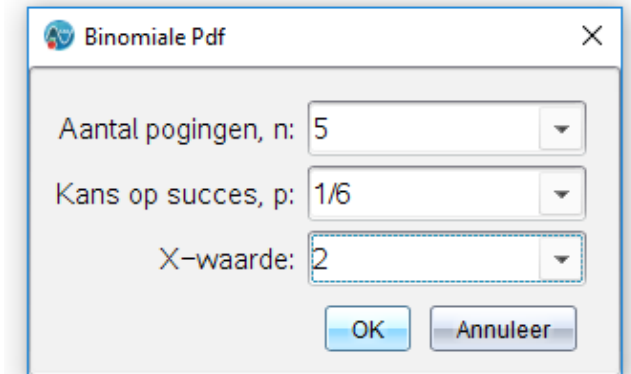
Stel dat in onze opleiding 10% meisjes zitten. Als je een groep van 50 studenten ziet, hoeveel meisjes verwacht je daar dan bij?

# Binomiale kansfunctie: rekentoestel

TI-Nspire: menu – 5: Kansen – 5: Verdelingen... – A: Binomiale Pdf...

Voorbeeld: 2x zes gooien bij 5 worpen met een dobbelsteen:

$$\text{binomPdf}\left(5, \frac{1}{6}, 2\right) = 0,160751$$



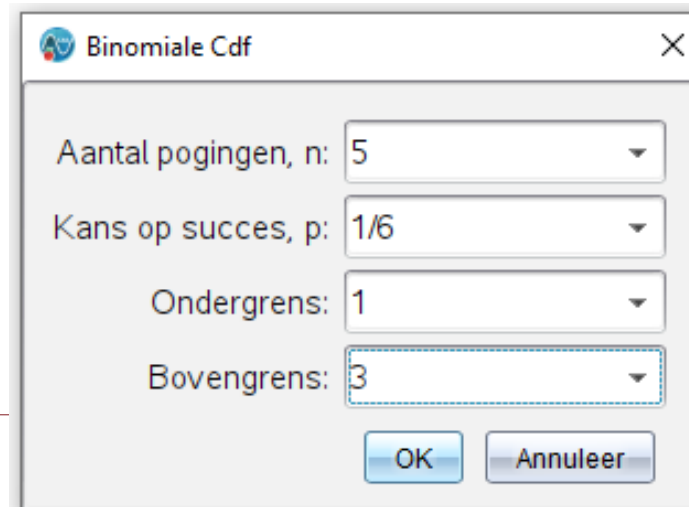
The screenshot shows the 'Binomiale Pdf' window on a TI-Nspire calculator. It has three input fields: 'Aantal pogingen, n:' with the value 5, 'Kans op succes, p:' with the value 1/6, and 'X-waarde:' with the value 2. At the bottom are 'OK' and 'Annuleer' buttons.

Of ook: menu – 5: Kansen – 5: Verdelingen... – B: Binomiale Cdf...

Voorbeeld: 1x of 2x of 3x een zes gooien bij 5 worpen met een dobbelsteen:

$$\text{binomCdf}\left(5, \frac{1}{6}, 1, 3\right) = 0,594779$$

interval



The screenshot shows the 'Binomiale Cdf' window on a TI-Nspire calculator. It has four input fields: 'Aantal pogingen, n:' with the value 5, 'Kans op succes, p:' with the value 1/6, 'Ondergrens:' with the value 1, and 'Bovengrens:' with the value 3. At the bottom are 'OK' and 'Annuleer' buttons.



# Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aardbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans  $P$  dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

A.  $0,05 \leq P < 0,1$

C.  $0,15 \leq P < 0,2$

B.  $0,1 \leq P < 0,15$

D.  $0,2 \leq P < 0,25$

# Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aardbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans  $P$  dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

Ga naar [wooclap.com](https://wooclap.com) en gebruik de code **STAT2802**



Hoe definieer je de binomiaal verdeelde kansvariabele?  $X = \dots$



wooclap



100 %



0



# Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aardbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans  $P$  dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

Ga naar [wooclap.com](https://wooclap.com) en gebruik de code **STAT2802**



Welke binomiale verdeling volgt de kansvariabele?



wooclap



100 %



0



# Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aardbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans  $P$  dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

Ga naar [wooclap.com](https://wooclap.com) en gebruik de code **STAT2802**



Schrijf de gezochte kans m.b.v. de kansvariabele  $X$ .



100 %



0



# Meerkeuzevraag

Bij Planet Yoghurt in Hasselt zijn aardbeien in de zomer een zeer populaire topping op frozen yoghurt (40% van de klanten kiest aardbeien). Op een dag zijn er te weinig aardbeien geleverd. In de namiddag ziet de uitbaatster dat er nog genoeg aardbeien zijn om 30 toppings te maken. Als er die dag nog 90 klanten komen, wat is dan de kans  $P$  dat elke liefhebber van aardbeien zijn favoriete topping kan kiezen?

U kunt niet langer stemmen



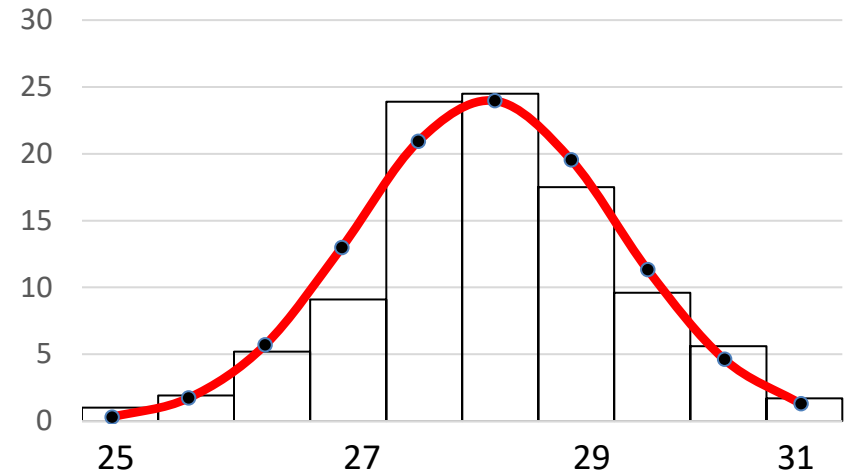
- | Nummer | Optie               | Percentage | Aantal Stemmen | Status |
|--------|---------------------|------------|----------------|--------|
| 1      | $0,05 \leq P < 0,1$ | 0%         | 0              |        |
| 2      | $0,1 \leq P < 0,15$ | 0%         | 0              | ✓      |
| 3      | $0,15 \leq P < 0,2$ | 0%         | 0              |        |
| 4      | $0,2 \leq P < 0,25$ | 0%         | 0              |        |

# Déjà vu...



Grootte van de voet

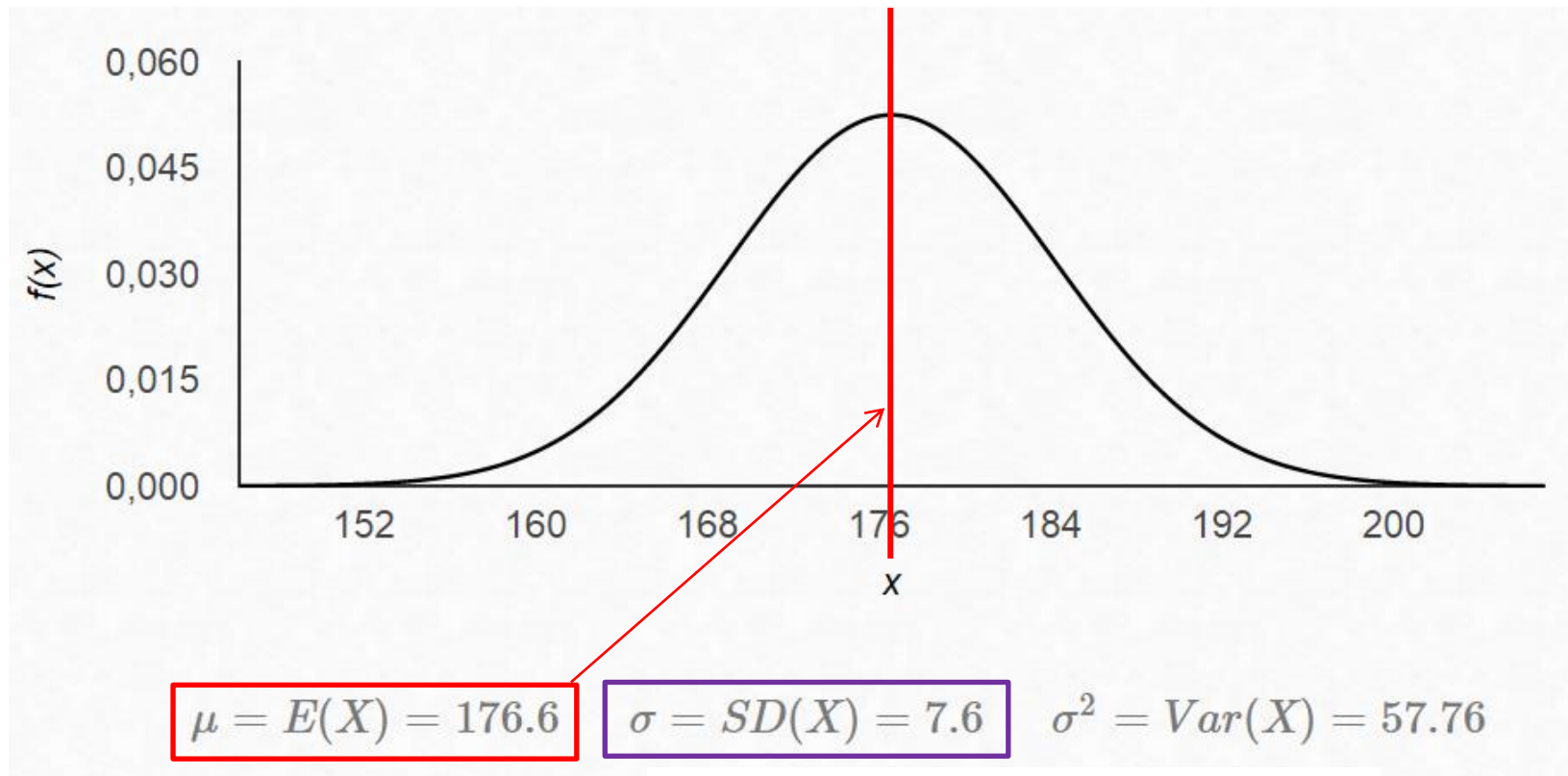
**CONTINUE  
KANSVARIABELE**



**CONTINUE  
KANSVERDELING**

# Normale kansdichtheid

Voorbeeld: de lengte van Belgische mannen (2005) heeft volgende kansdichtheid:



$$X \sim N(\mu, \sigma)$$

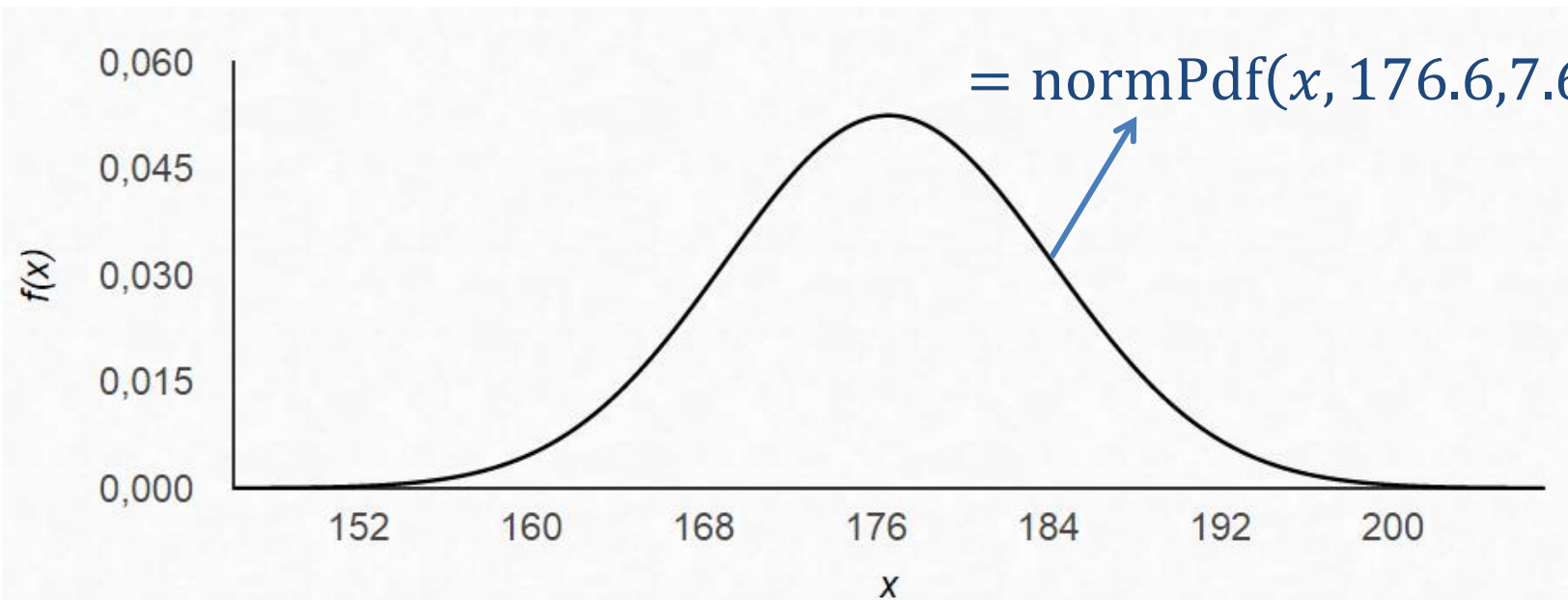
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Hoe kleiner  $\sigma$ , hoe kleiner de spreiding, hoe scherper de piek

# Normale kansdichtheid

Voorbeeld: de lengte van Belgische mannen (2005) heeft volgende kansdichtheid:

Continue kansvariabele



$= \text{normPdf}(x, 176.6, 7.6) = \text{hoogte kansdichtheid}$   
 $\neq$  kans

vs.  $\text{binomPdf} = \text{kans}$

Discrete kansvariabele

$$\mu = E(X) = 176.6 \quad \sigma = SD(X) = 7.6 \quad \sigma^2 = Var(X) = 57.76$$



# Normale kansdichtheid

The image shows a software interface for a document editor. On the left is a TI-Nspire CX II-T CAS calculator. The main area is a document editor with a toolbar at the top. The document content shows the formula  $\text{normPdf}(170, 176.6, 7.6)$  and the result  $0.036003$ .

Documentenwerkset

Toetsenbord + Zijscherm

TI-Nspire CX II-T CAS

esc save tab ctrl caps shift var clear del trig 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

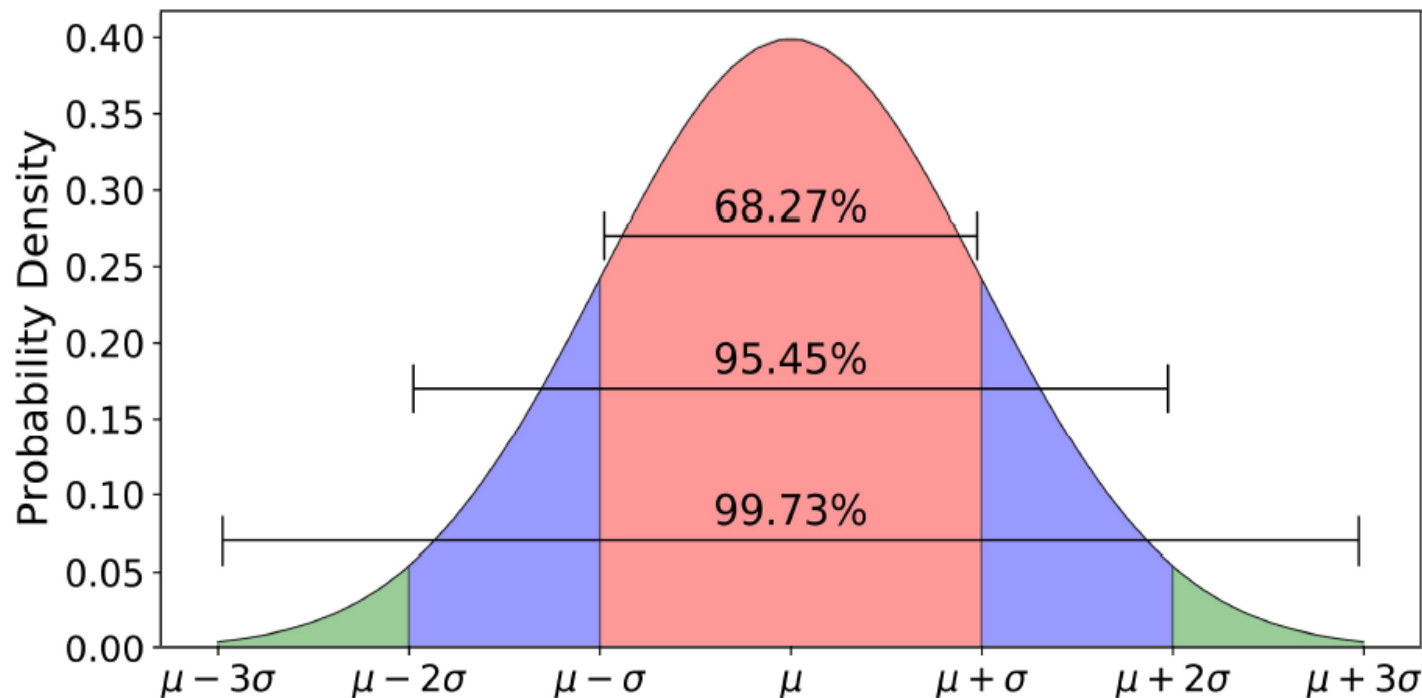
normPdf(170,176.6,7.6)

0.036003

# Algemene normale kansdichtheid: eigenschappen

**Eigenschap 6.6** *Een normaal verdeelde variabele neemt*

- *in 68.3% van de gevallen een waarde aan tussen  $\mu - \sigma$  en  $\mu + \sigma$ .*
- *in 95.5% van de gevallen een waarde aan tussen  $\mu - 2\sigma$  en  $\mu + 2\sigma$ .*
- *in 99.7% van de gevallen een waarde aan tussen  $\mu - 3\sigma$  en  $\mu + 3\sigma$ .*



Reken zelf na!

`normCdf(128,132,130,2) = ...`

`normCdf(75,95,85,5) = ...`

`normCdf(750,1650,1200,150) = ...`

# Normale verdeling: rekentoestel

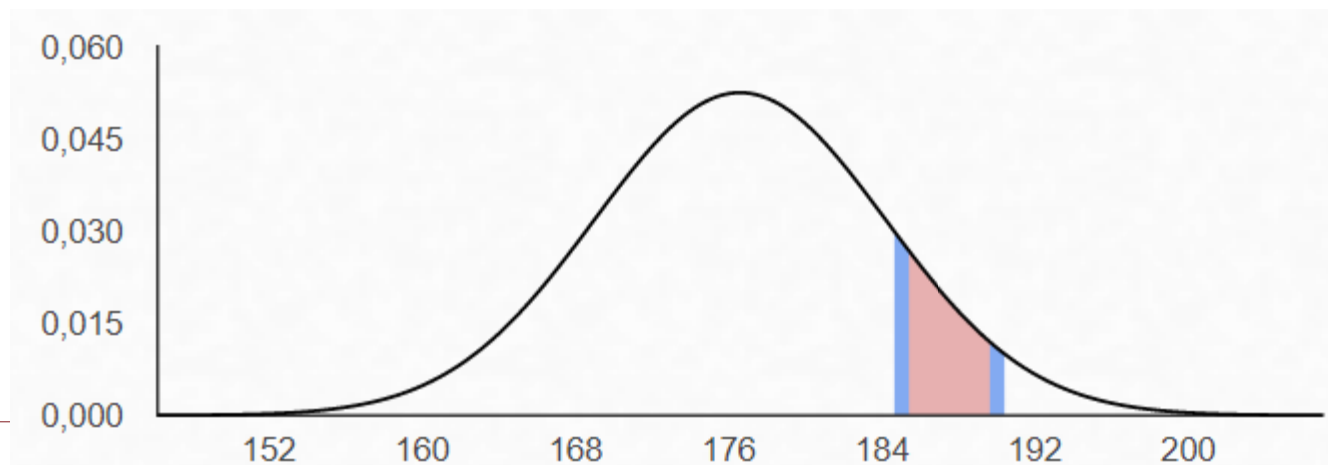
- menu – 6: Statistieken – 5: Verdelingen... – 1: Normale Pdf...

Berekent de functiewaarde van de dichtheidsfunctie (zullen we bijna niet gebruiken)

- menu – 5: Kansen – 5: Verdelingen... – 2: Normale Cdf...

Voorbeeld: Bereken de kans dat de lengte van een Belgische man tussen 185 cm en 190 cm ligt.

$$P(185 < X < 190) = \text{normCdf}(185, 190, 176.6, 7.6) = 0.095586$$

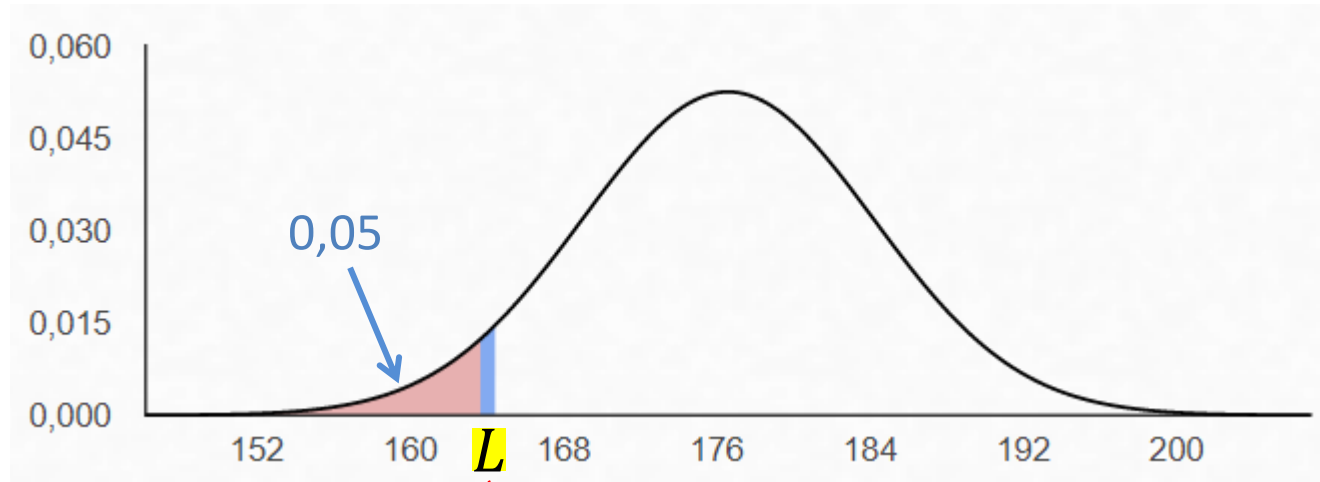


# Normale verdeling: rekentoestel

- menu – 6: Statistieken – 5: Verdelingen... – 3: **Inverse-normaal...**

Voorbeeld: Bij welke lengte ligt het 5% percentiel (of het 0,05-kwantiel) van de Belgische mannen?

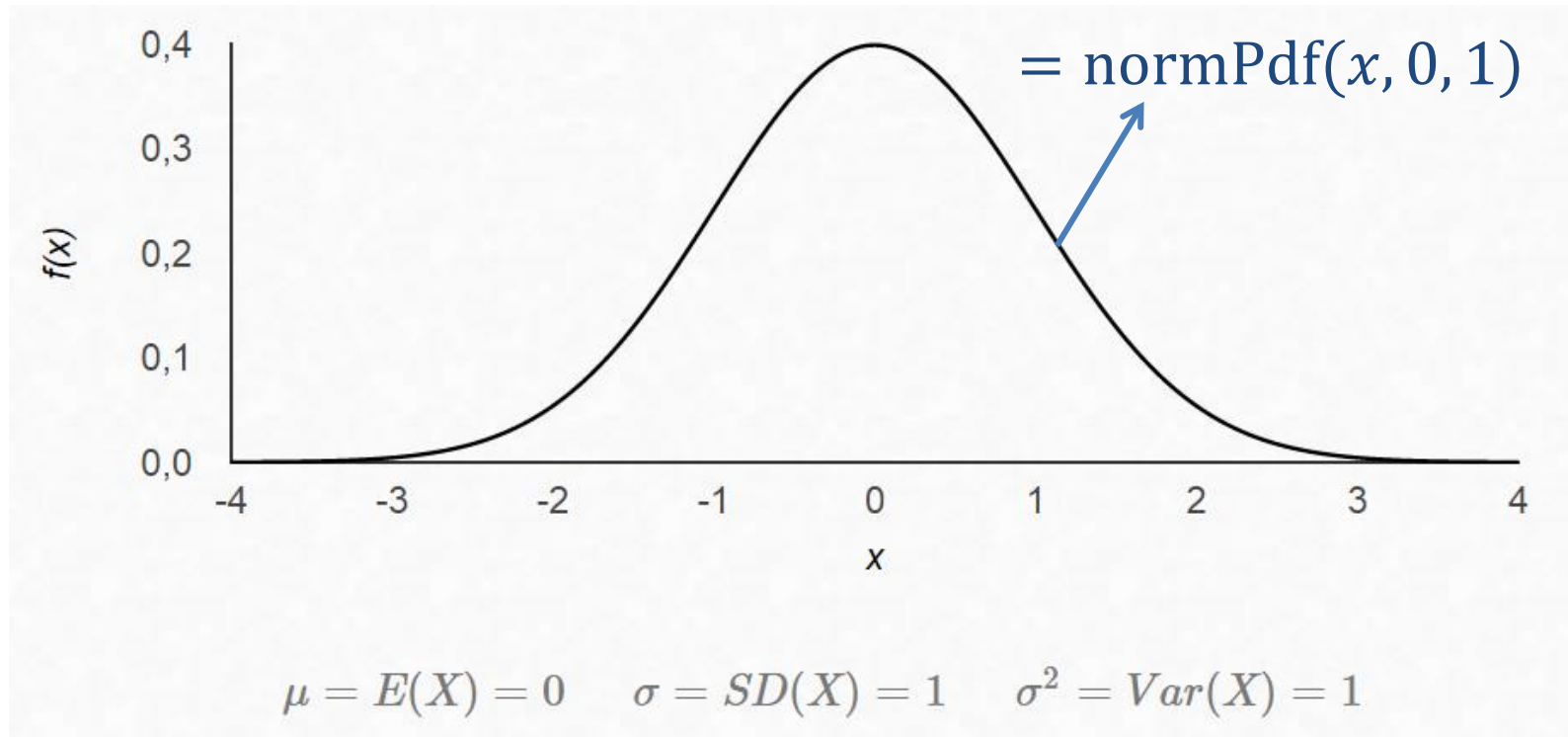
$$P(X < L) = 0.05 \Rightarrow L = \text{invNorm}(0.05, 176.6, 7.6) = 164.099 \text{ cm}$$



waarde op horizontale as zoeken

A screenshot of a software window titled 'Inverse-normaal'. It contains three input fields: 'Oppervlakte:' with a dropdown menu showing '0.05', 'μ:' with a dropdown menu showing '176.6', and 'σ:' with a dropdown menu showing '7.6'. Below these fields are two buttons: 'OK' and 'Annuleer'.

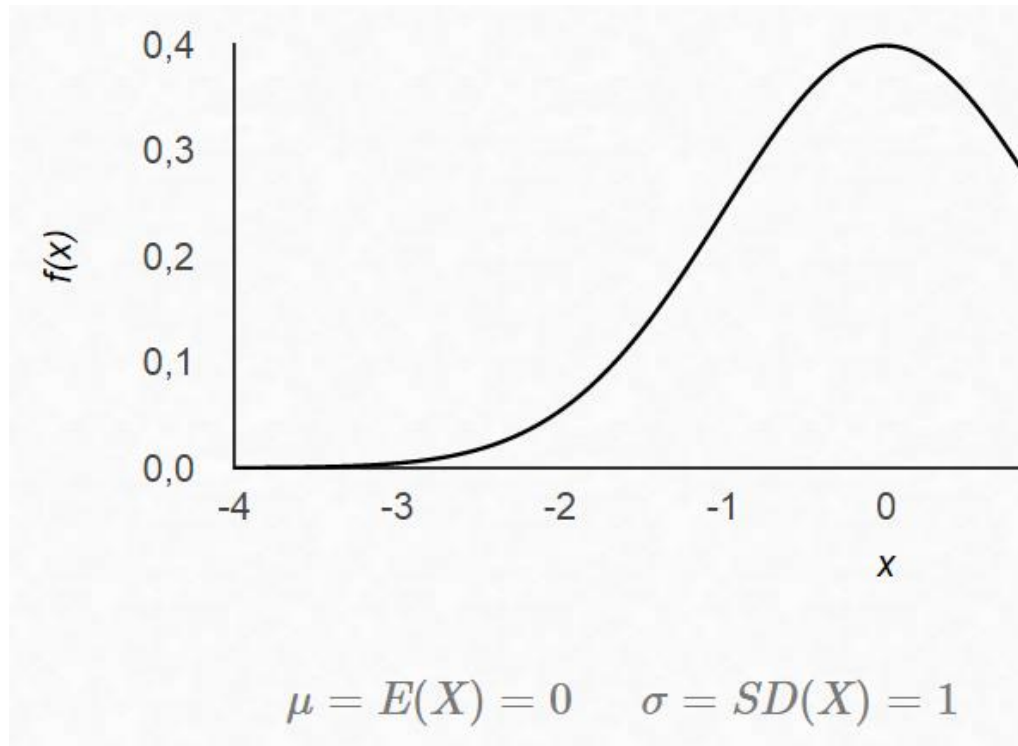
# De standaard normale kansdichtheid



$$Z \sim N(0, 1)$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

# De standaard normale kansdichtheid



$= \text{normPdf}(x, 0, 1)$



$Z \sim N(0, 1)$

$z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8341
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8979
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292

# Gestandaardiseerde kansvariabele

**Eigenschap 6.4** Als  $X \sim N(\mu, \sigma)$ , dan heeft de gestandaardiseerde kansvariabele

$$Z = \frac{X - \mu}{\sigma}$$

een standaard normale kansdichtheid  $N(0, 1)$  met verwachting 0 en standaardafwijking 1.

**LINEAIRE COMBINATIES:** Als  $X_1, X_2, \dots, X_n$  onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting  $\mu_i$  en standaardafwijking  $\sigma_i$  voor  $i = 1, 2, \dots, n$  dan is  $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$  ook normaal verdeeld met verwachting  $\mu = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n + b$  en standaardafwijking  $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2}$ . (opmerking: hierbij mogen de constanten  $a_i$  en  $b$  negatief zijn !!)

# Oef 39 p17 – zie werksitting

De levensduur van een machine is normaal verdeeld met een verwachtingswaarde van 3000 uren. Uit ervaring is gebleken dat 50% van deze machines minder dan 2632 uren of meer dan 3368 uren werken. Bereken de standaardafwijking van de levensduur van deze machines.



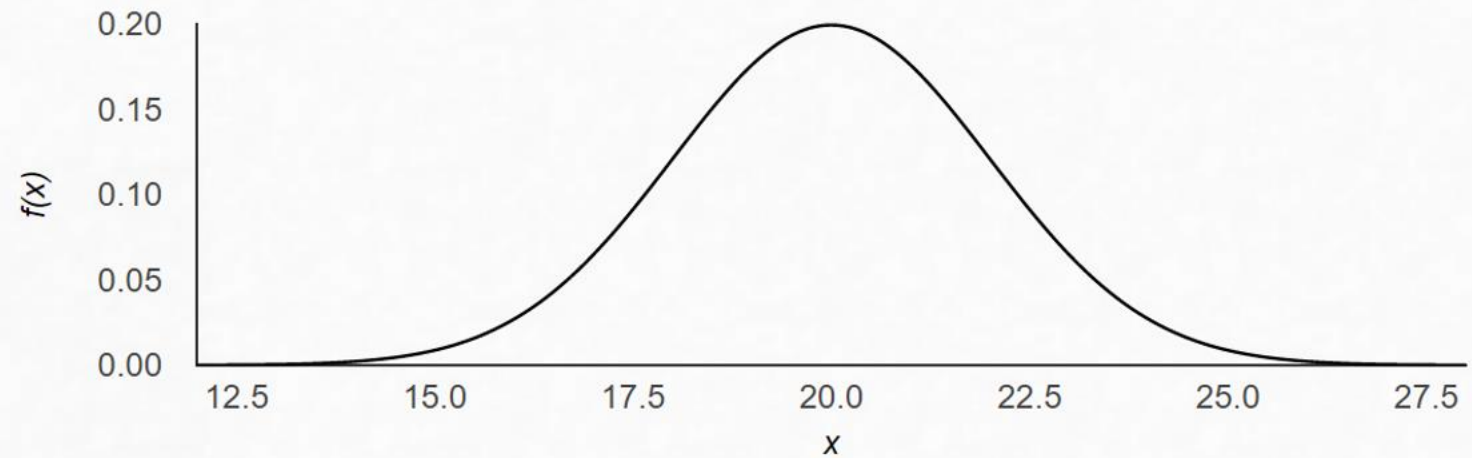
# Meerkeuzevraag

U kunt niet langer stemmen

Bij welke normale verdeling hoort deze grafiek?

- 1  $N(20, 2)$  0% 0 ✓
- 2  $N(20, 5)$  0% 0
- 3  $N(20, 7.5)$  0% 0
- 4  $N(15, 2)$  0% 0
- 5  $N(15, 5)$  0% 0
- 6  $N(15, 7.5)$  0% 0

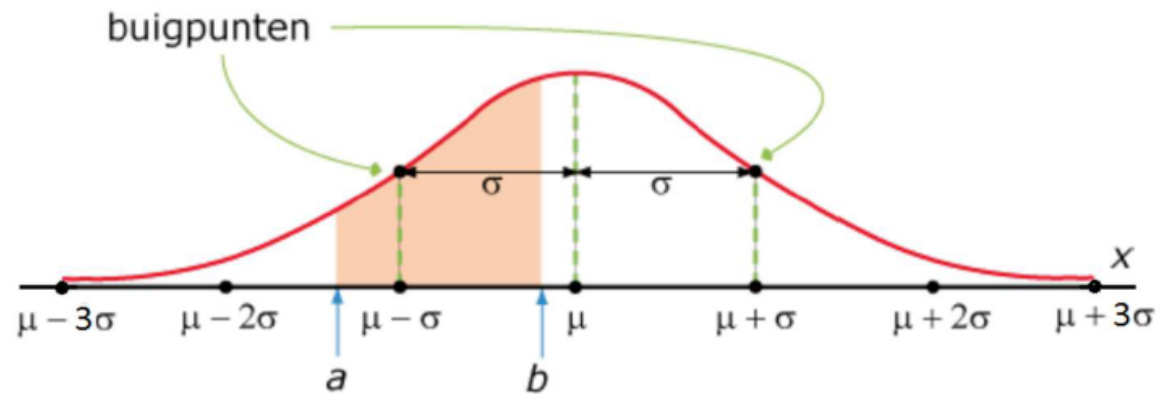
Bij welke normale verdeling hoort deze grafiek?



# Formularium

KANSDICHTHEIDSFUNCTIE:  $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$  ← rekentoestel: normPdf(x, μ, σ)

NOTATIE:  $X \sim N(\mu, \sigma) \rightarrow E(X) = \mu, \text{Var}(X) = \sigma^2$



KANSVERDELINGSFUNCTIE:  $P(X \leq x) = \int_{-\infty}^x f_X(t) dt$  ← rekentoestel: normCdf(-∞, x, μ, σ)

KANS MET ONDER- en BOVENGRENS:  $P(a \leq X \leq b) = \text{normCdf}(a, b, \mu, \sigma)$

**BELANGRIJKSTE  $\sigma$ -NIVEAUS:** Voor elke normale kansverdeling  $X \sim N(\mu, \sigma)$  geldt

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 68,27\%$$

$$P(\mu - 1.645\sigma \leq X \leq \mu + 1.645\sigma) \approx 90\%$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 95,45\%$$

$$P(\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma) \approx 95\%$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 99,73\%$$

$$P(\mu - 2.576\sigma \leq X \leq \mu + 2.576\sigma) \approx 99\%$$

**p-KWANTIELEN** ( $0 \leq p \leq 1$ ): zoek  $c$  zodat  $P(X \leq c) = p$  ← rekentoestel: invNorm(p, μ, σ)

**LINK met STANDAARDNORMALE VERDELING:** Als  $X \sim N(\mu, \sigma)$ , dan heeft de gestandaardiseerde kansvariabele  $Z = \frac{X - \mu}{\sigma}$  (=Z-score) een normale verdeling  $N(0,1)$  met verwachting 0 en standaardafwijking 1.

# De Poissonverdeling – voorbeeld

In de kerkstraat in Diepenbeek worden regelmatig snelheidscontroles gedaan. Gemiddeld worden er 28 auto's per uur geflitst. Wat is de kans dat er bij de volgende controle 35 auto's per uur geflitst worden?

# De Poissonverdeling – voorbeeld

In de kerkstraat in Diepenbeek worden regelmatig snelheidscontroles gedaan. Het aantal auto's dat geflitst wordt is Poisson verdeeld met een gemiddelde van 28 auto's per uur. Wat is de kans dat er bij de volgende controle 35 auto's per uur geflitst worden?

# Poisson: voorwaarden

1.  $X$  is het aantal gebeurtenissen in een continu interval van een bepaalde vaste grootte.
2. Het verwachte aantal gebeurtenissen in 2 intervallen van dezelfde lengte is gelijk en is evenredig met de lengte.
3. Het aantal gebeurtenissen in een interval is onafhankelijk van het aantal gebeurtenissen in een niet-overlappend interval.

Dit is niet zo gemakkelijk als het lijkt!

Voorbeeld: bakker maakt krentenbrood met gemiddeld 8 krenten per snee brood.

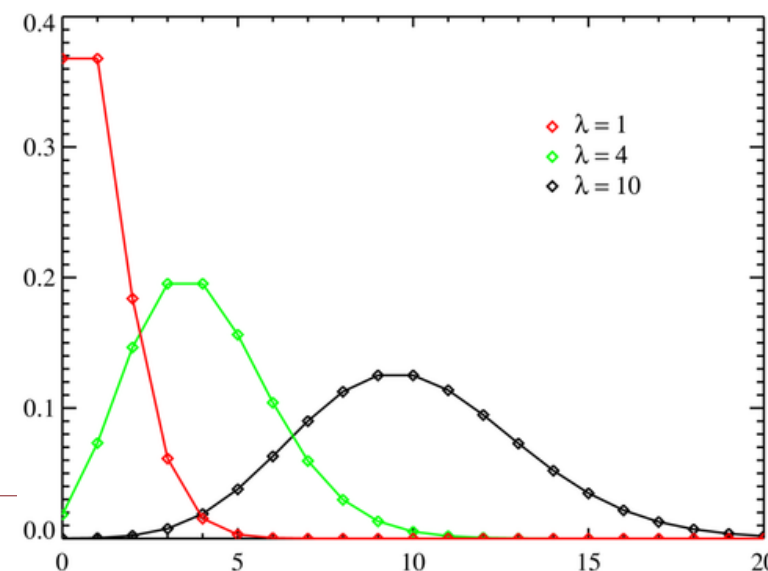
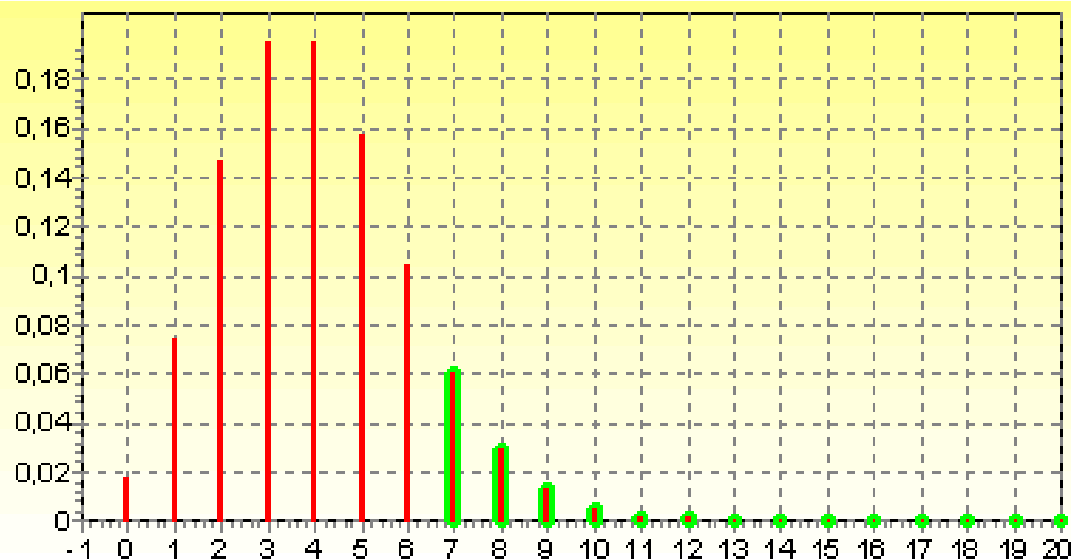
5	4	9	5	10	13	10	8
---	---	---	---	----	----	----	---

# Poisson: Formularium

Noem  $X$  het aantal gebeurtenissen in een continu interval (bvb. een bepaalde afstand of een tijdsinterval).  
Wanneer het **verwachte** aantal gebeurtenissen  $\lambda$  **evenredig** is met de lengte van het interval en bovendien het aantal gebeurtenissen die optreden in 2 niet-overlappende (deel)intervallen *onafhankelijk* zijn van elkaar, dan is de kansvariabele  $X$  **Poisson** verdeeld met gemiddelde  $\lambda$ :  $X \sim Poi(\lambda)$ .

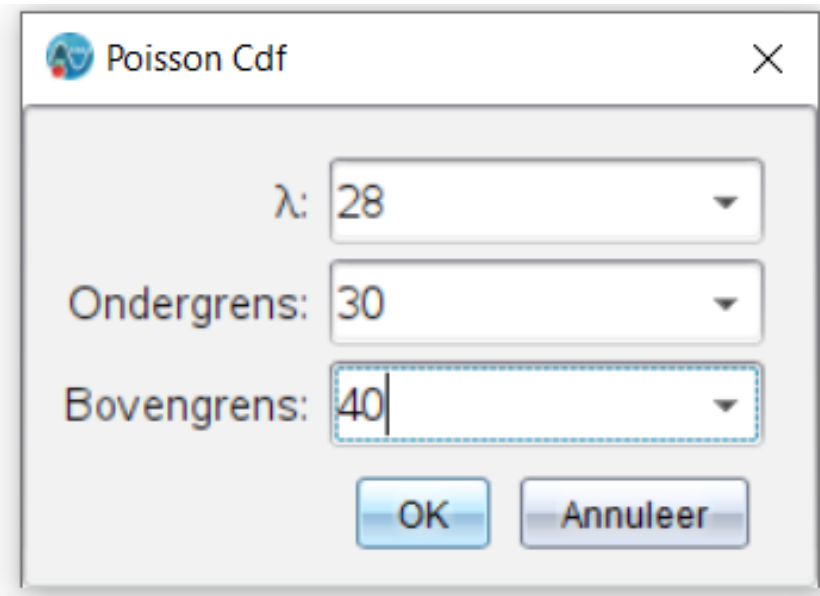
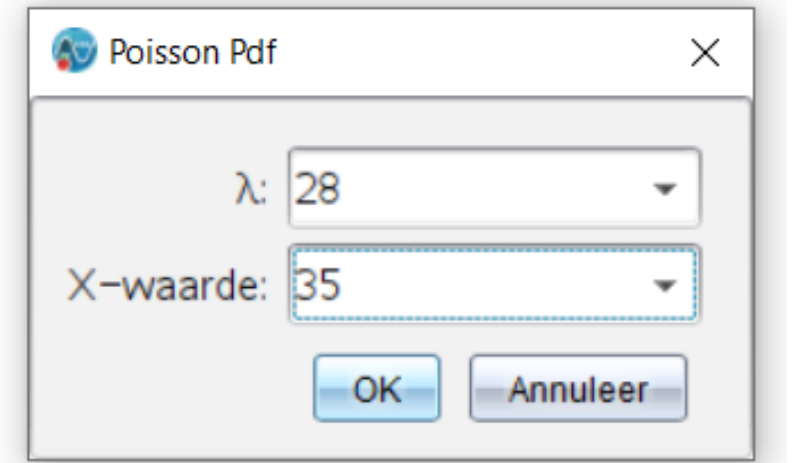
De bijhorende kansfunctie:  $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$  ← rekentoestel: `poissPdf( $\lambda, k$ )`

Voor een Poisson verdeling geldt:  $\mu_X = \lambda$  en  $\sigma_X = \sqrt{\lambda}$ .



# Poisson: rekentoestel

- Notatie:  $X \sim Poi(\lambda)$
- TI-Nspire: menu – 6: Statistieken – 5: Verdelingen... – H: Poisson Pdf...  
of I: Poisson Cdf...



# Voorbeeld: snelheidscontrole

In de kerkstraat in Diepenbeek worden regelmatig snelheidscontroles gedaan. Het aantal auto's dat geflitst wordt is Poisson verdeeld met een gemiddelde van 28 auto's per uur. Wat is de kans dat er bij de volgende controle 35 auto's per uur geflitst worden?

$$P(35 \text{ auto's}) = \text{poissPdf}(28, 35) = 0.029926$$

Wat is de kans dat er tussen de 30 en de 40 auto's per uur geflitst worden?

$$P(30 \text{ tot } 40 \text{ auto's}) = \text{poissCdf}(28, 30, 40) = 0.364925$$



# Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans  $P$  dat er meer dan 25 bezoekers zijn in 5 minuten?

A.  $0.05 \leq P \leq 0.10$

B.  $0.10 \leq P \leq 0.15$

C.  $0.15 \leq P \leq 0.20$

D.  $0.20 \leq P \leq 0.25$

# Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans  $P$  dat er meer dan 25 bezoekers zijn in 5 minuten?

Ga naar **wooclap.com** en gebruik de code **STAT2802**



Hoe definieer je de kansvariabele bij deze vraag?  $X = \dots$



# Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans  $P$  dat er meer dan 25 bezoekers zijn in 5 minuten?

Ga naar **wooclap.com** en gebruik de code **STAT2802**



Welke kansverdeling volgt de kansvariabele?



# Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans  $P$  dat er meer dan 25 bezoekers zijn in 5 minuten?

Ga naar [wooclap.com](https://wooclap.com) en gebruik de code **STAT2802**



Schrijf de gezochte kans m.b.v. de kansvariabele  $X$ .



# Meerkeuzevraag

Het aantal bezoekers per minuut op een webpagina is Poisson verdeeld. Als het gemiddeld aantal bezoekers per minuut gelijk is aan 4, wat is dan de kans  $P$  dat er meer dan 25 bezoekers zijn in 5 minuten?

U kunt niet langer stemmen



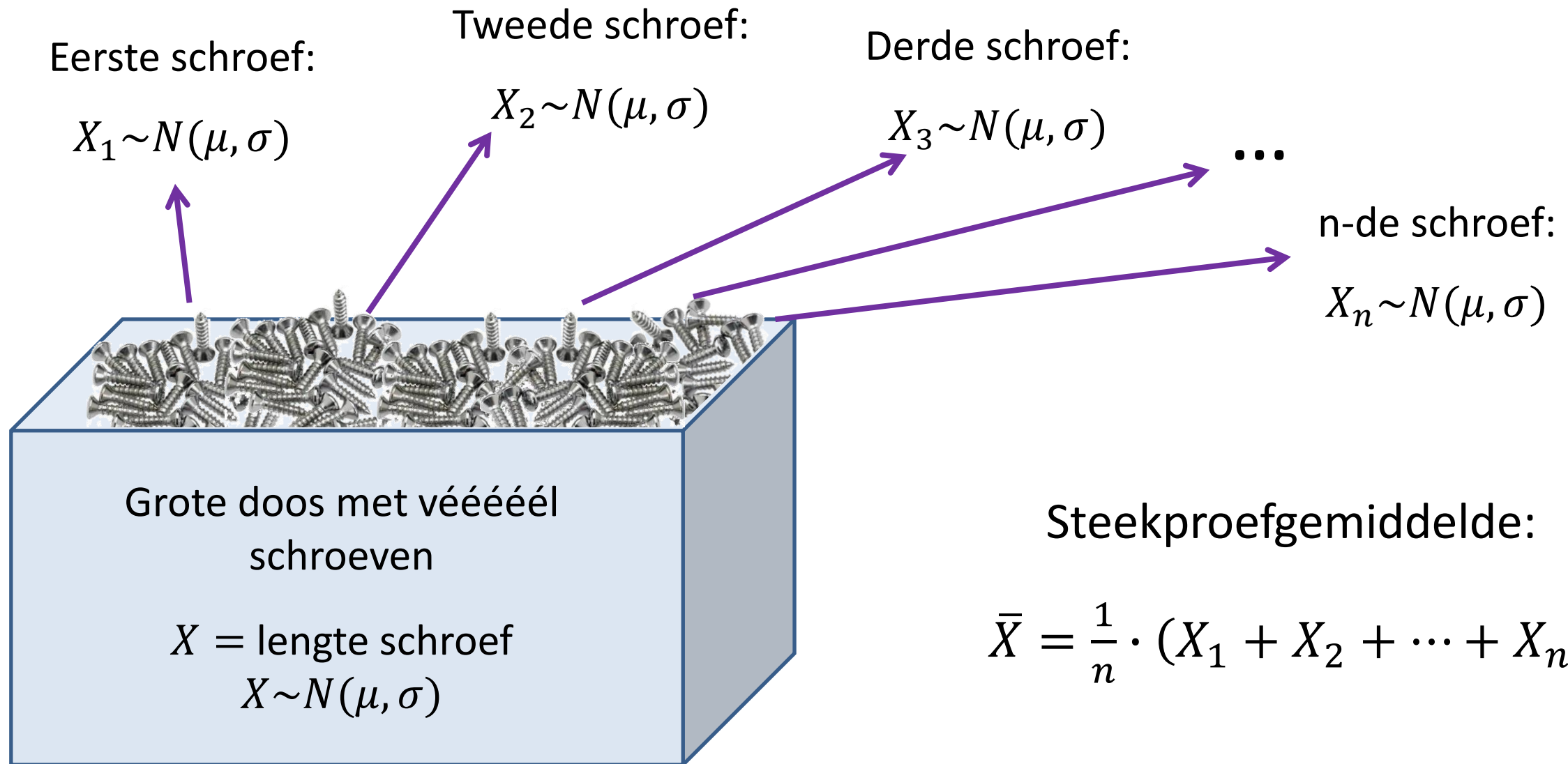
En nu nog berekenen! In welk interval ligt de gevraagde kans  $P$ ?

- 1  $0.05 \leq P \leq 0.10$  0% 0
- 2  $0.10 \leq P \leq 0.15$  0% 0 ✓
- 3  $0.15 \leq P \leq 0.20$  0% 0
- 4  $0.20 \leq P \leq 0.25$  0% 0

## Déjà vu: formularium

**LINEAIRE COMBINATIES:** Als  $X_1, X_2, \dots, X_n$  onderling **onafhankelijke** ~~normaal verdeelde~~ toevalsvariabelen zijn met verwachting  $\mu_i$  en standaardafwijking  $\sigma_i$  voor  $i = 1, 2, \dots, n$  dan is  $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$  ook ~~normaal verdeeld~~ met verwachting  $\mu = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n + b$  en standaardafwijking  $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2}$ . (opmerking: hierbij mogen de constanten  $a_i$  en  $b$  negatief zijn !!)

# Lineaire combinaties: (steekproef)gemiddelde



# Lineaire combinaties: (steekproef)gemiddelde

**Gevolg 6.9** Als  $X_i \sim N(\mu, \sigma)$ ,  $i = 1, \dots, n$  onderling onafhankelijk, dan is

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

ook normaal verdeeld met gemiddelde  $\mu$  en standaardafwijking  $\frac{\sigma}{\sqrt{n}}$ .

$$\mu_{\bar{X}} = \frac{1}{n} \cdot \mu_{X_1} + \frac{1}{n} \cdot \mu_{X_2} + \dots + \frac{1}{n} \cdot \mu_{X_n} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

**LINEAIRE COMBINATIES:** Als  $X_1, X_2, \dots, X_n$  onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting  $\mu_i$  en standaardafwijking  $\sigma_i$  voor  $i = 1, 2, \dots, n$  dan is  $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$  ook normaal verdeeld met verwachting  $\mu = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n + b$  en standaardafwijking  $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2}$ . (opmerking: hierbij mogen de constanten  $a_i$  en  $b$  negatief zijn !!)



# Lineaire combinaties: (steekproef)gemiddelde

**Gevolg 6.9** Als  $X_i \sim N(\mu, \sigma)$ ,  $i = 1, \dots, n$  onderling onafhankelijk, dan is

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

ook normaal verdeeld met gemiddelde  $\mu$  en standaardafwijking  $\frac{\sigma}{\sqrt{n}}$ .

$$\mu_{\bar{X}} = \frac{1}{n} \cdot \mu_{X_1} + \frac{1}{n} \cdot \mu_{X_2} + \dots + \frac{1}{n} \cdot \mu_{X_n} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

$$\mu_{\bar{X}} = \mu$$

Het **gemiddelde** van het **steekproefgemiddelde** is gelijk aan het **populatiegemiddelde**

# Lineaire combinaties: (steekproef)gemiddelde

**Gevolg 6.9** Als  $X_i \sim N(\mu, \sigma)$ ,  $i = 1, \dots, n$  onderling onafhankelijk, dan is

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

ook normaal verdeeld met gemiddelde  $\mu$  en standaardafwijking  $\frac{\sigma}{\sqrt{n}}$ .

$$\mu_{\bar{X}} = \frac{1}{n} \cdot \mu_{X_1} + \frac{1}{n} \cdot \mu_{X_2} + \dots + \frac{1}{n} \cdot \mu_{X_n} = \frac{1}{n} \cdot n \cdot \mu = \mu$$

$$\sigma_{\bar{X}} = \sqrt{\frac{1}{n^2} \cdot \sigma_{X_1}^2 + \frac{1}{n^2} \cdot \sigma_{X_2}^2 + \dots + \frac{1}{n^2} \cdot \sigma_{X_n}^2} = \sqrt{\frac{1}{n^2} \cdot n \cdot \sigma^2} = \sqrt{\frac{\sigma^2}{n}}$$

**LINEAIRE COMBINATIES:** Als  $X_1, X_2, \dots, X_n$  onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting  $\mu_i$  en standaardafwijking  $\sigma_i$  voor  $i = 1, 2, \dots, n$  dan is  $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$  ook normaal verdeeld met verwachting  $\mu = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n + b$  en standaardafwijking  $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2}$ . (opmerking: hierbij mogen de constanten  $a_i$  en  $b$  negatief zijn !!)

# Lineaire combinaties: (steekproef)gemiddelde

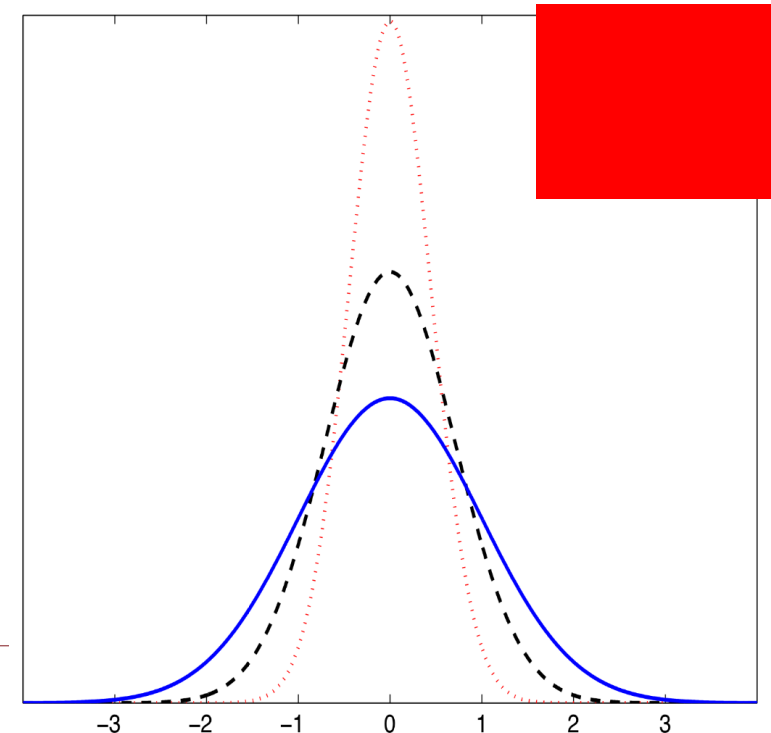
Een lineaire combinatie die veelvuldig voorkomt is het gemiddelde van  $n$  onderling onafhankelijke normaal verdeelde variabelen met dezelfde verwachting  $\mu$  en standaardafwijking  $\sigma$ . Dan is het gemiddelde ook normaal verdeeld met verwachting  $\mu$  en kleinere standaardafwijking  $\sigma/\sqrt{n}$ . Dus

**Gevolg 6.9** Als  $X_i \sim N(\mu, \sigma)$ ,  $i = 1, \dots, n$  onderling onafhankelijk, dan is

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

ook normaal verdeeld met gemiddelde  $\mu$  en standaardafwijking  $\frac{\sigma}{\sqrt{n}}$ .

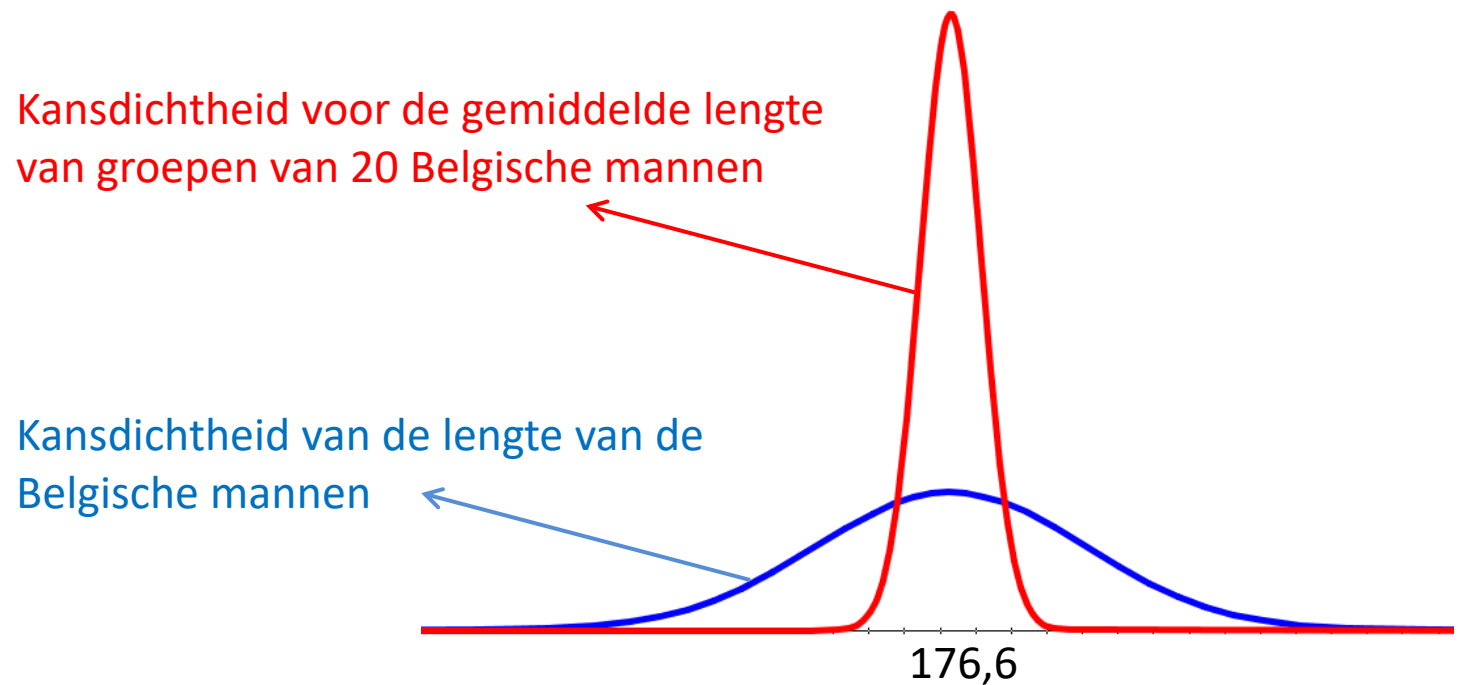
Hiernaast zie je de kansverdeling van 3 steekproeven.  
Welke verdeling hoort bij de grootste steekproef?



# Lineaire combinaties: (steekproef)gemiddelde

Voorbeeld: de gemiddelde lengte  $\bar{X}$  van een groep van 20 Belgische mannen

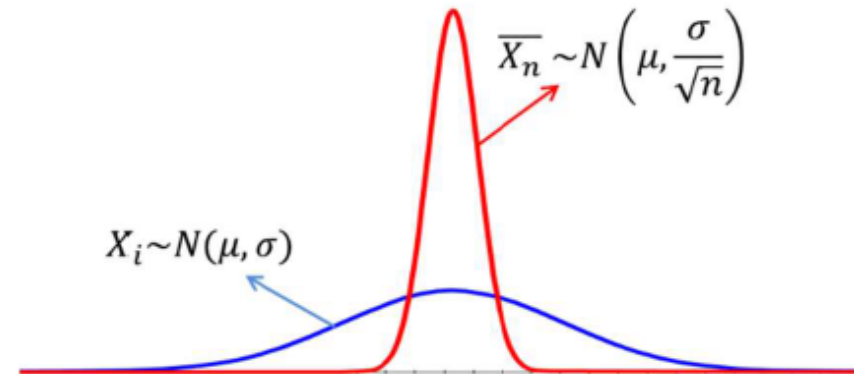
$$\bar{X} \sim N(176.6, \frac{7.6}{\sqrt{20}})$$



# Formularium

**LINEAIRE COMBINATIES:** Als  $X_1, X_2, \dots, X_n$  onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting  $\mu_i$  en standaardafwijking  $\sigma_i$  voor  $i = 1, 2, \dots, n$  dan is  $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n + b$  ook normaal verdeeld met verwachting  $\mu = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n + b$  en standaardafwijking  $\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2}$ . (opmerking: hierbij mogen de constanten  $a_i$  en  $b$  negatief zijn !!)

**STEEKPROEFGEMIDDELDE:** Als  $X_1, X_2, \dots, X_n$  onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting  $\mu$  en standaardafwijking  $\sigma$  voor  $i = 1, 2, \dots, n$  dan is  $\bar{X}_n \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ .



# Benaderende verdelingen

Binomale verdeling:

- Poisson benadering
- Normale benadering

Poisson verdeling:

- Normale benadering

Dit wordt nuttig als we hypothesetoetsen gaan uitvoeren

# Poisson benadering van de binomiale verdeling

## 7.1.4 Poisson benadering van binomiale verdeling

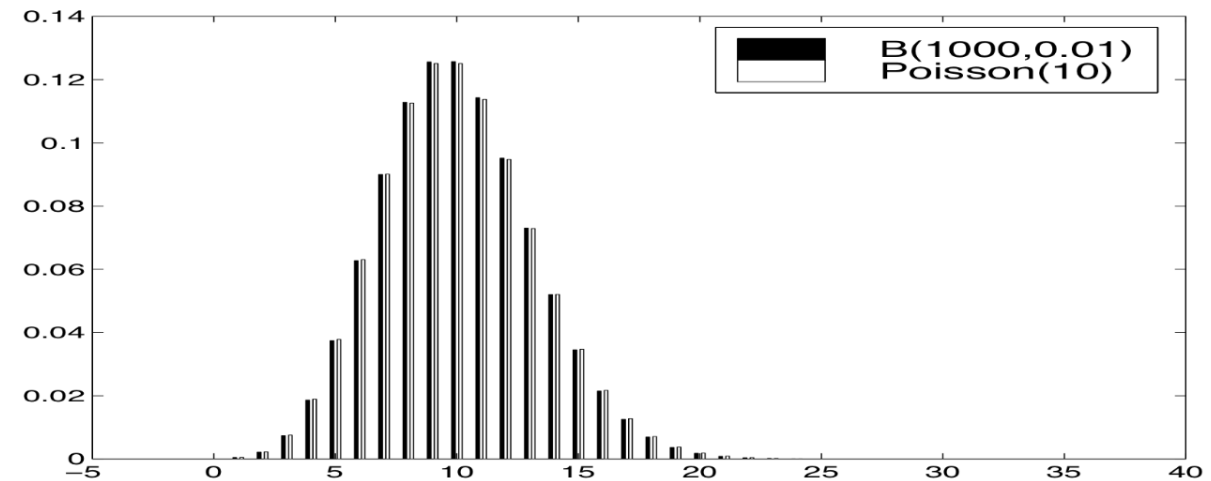
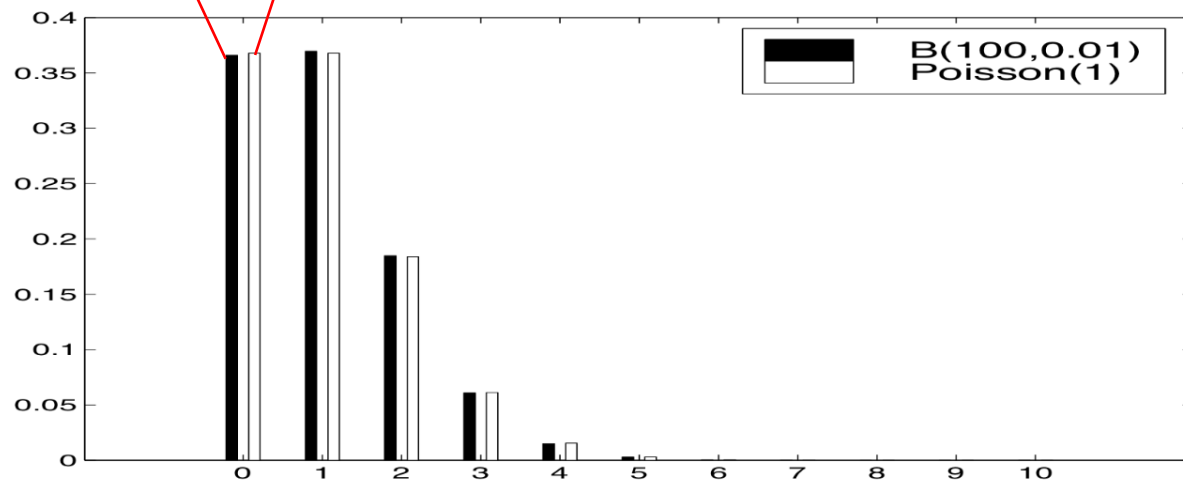
**Stelling 7.4** Als  $X \sim B(n, p)$  met het aantal deexperimenten  $n$  groot en de kans op succes  $p$  klein, dan is  $X$  bij benadering Poisson verdeeld met parameter  $\lambda = np$ . → Gemiddelde  $B(n, p)$

We zeggen dat  $n$  groot is en  $p$  klein als  $n \geq 50$  en  $p \leq 0.1$ . Dan is de benadering voldoende goed. Zie figuur 7.1 voor een illustratie.

binomPdf(100,0.01,0)

poissPdf(1,0)

Figuur 7.1: Illustratie van de Poisson benadering van de binomiale verdeling.



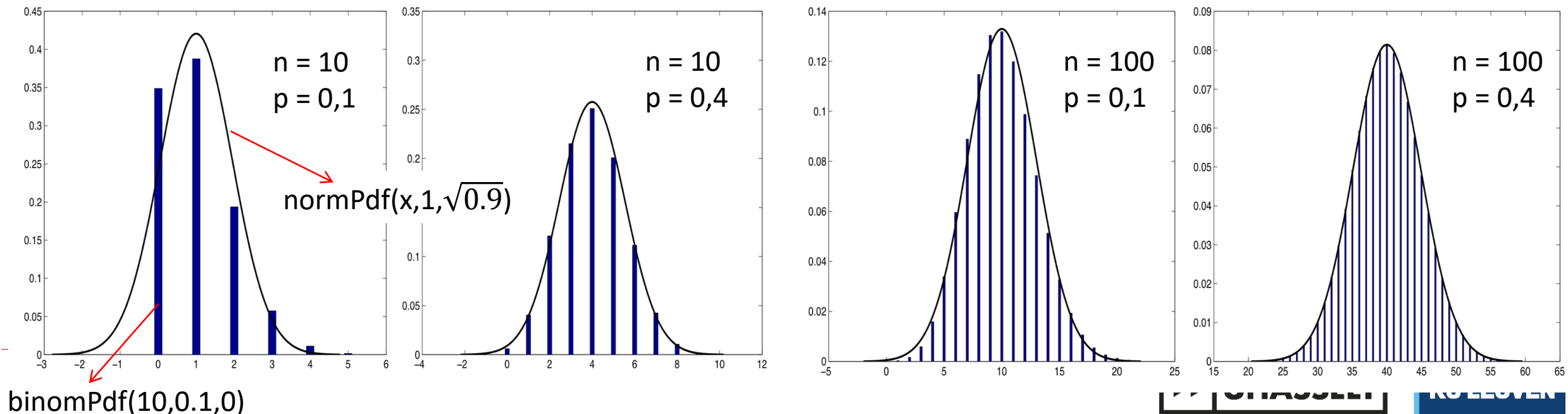
# Normale benadering van de binomiale verdeling

**Eigenschap 6.10** Als  $X \sim B(n, p)$  dan is  $X$  bij benadering

$$X \sim N\left(np, \sqrt{np(1-p)}\right)$$

verdeeld voor  $n$  groot.

Voor vaste  $n$  is deze benadering het beste voor  $p$  dicht bij 0.5. In praktijk is het een goede benadering als  $np \geq 5$  en  $n(1-p) \geq 5$ . Zie figuur 6.3 voor een illustratie.





# Normale benadering van de poisson verdeling

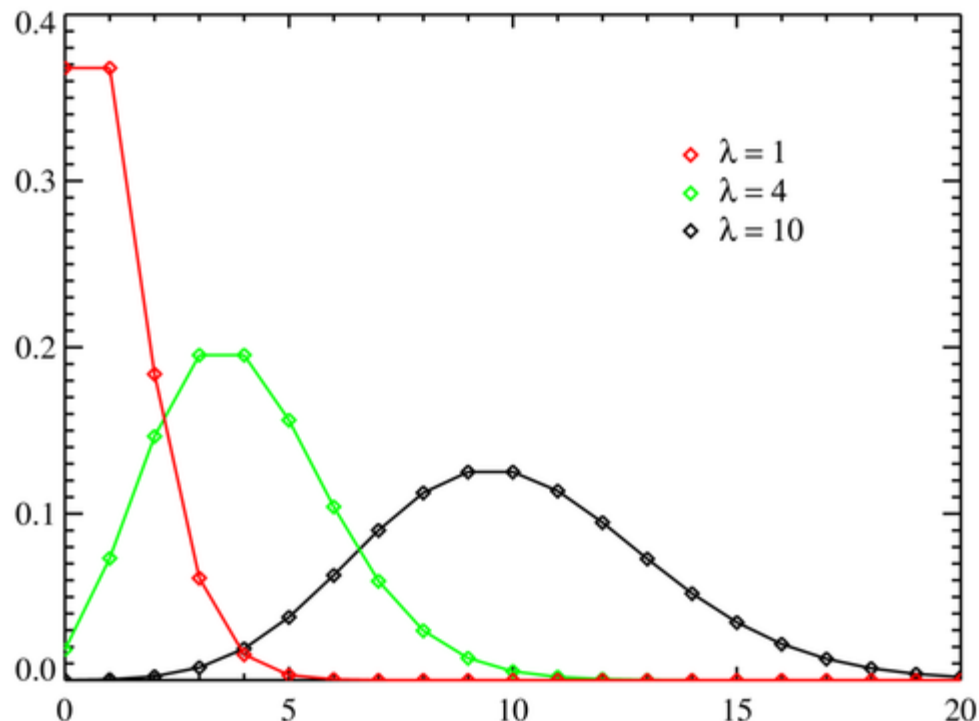
**Eigenschap 7.5** Als  $X$  Poisson verdeeld is met parameter  $\lambda$  en  $\lambda$  is groot, dan is  $X$  bij benadering normaal verdeeld:

$$X \sim N(\lambda, \sqrt{\lambda})$$

In de praktijk is dit een goede benadering als  $\lambda \geq 10$ .

$\mu$

$\sigma$



# Formularium

**NORMALE BENADERING:** Als  $X \sim B(n, p)$ , dan is  $X$  bij benadering  $N(np, \sqrt{np(1-p)})$  als  $np > 5$  en  $n(1-p) > 5$ .

Als  $X \sim \text{Poi}(\lambda)$ , dan is  $X$  bij benadering  $N(\lambda, \sqrt{\lambda})$  als  $\lambda > 10$ .

# TODO vóór de werksitting

Oefeningen 32 + 34 bekijken (oplossing staat op Toledo).  
Dit duurt ongeveer 10 minuten per oefening.