Simulatie gemiddelde bij verschillende verdelingen

Gebruik de app die je vindt op:

http://onlinestatbook.com/stat\_sim/sampling\_dist/index.html.

Net als bij de dobbelstenen, bestuderen we met deze app ook de verdeling van het gemiddelde van stochasten. In plaats van kansen te berekenen, werkt deze app met simulaties. Dit hadden we met de dobbelstenen ook kunnen doen: om de verdeling van het gemiddelde aantal ogen bij drie dobbelstenen te benaderen, hadden we een groot aantal keren drie dobbelstenen kunnen gooien en telkens het gemiddelde aantal ogen berekenen. Het histogram van die resultaten zou dan een benadering vormen voor het histogram van de theoretische kansen uit de vorige werktekst.

De bovenste verdeling kun je instellen. Activeer de knop ‘begin’ en selecteer als verdeling ‘uniform’. Dan zitten we in dezelfde situatie als bij de dobbelstenen: elke waarde van de stochasten heeft dezelfde kans.

1. Wat is de betekenis van wat er links van de uniforme verdeling staat?

Bij de derde grafiek staat ‘mean’ (gemiddelde) en ‘N=5’. Dit betekent dat we het gemiddelde nemen van 5 stochasten met de verdeling uit de eerste grafiek. Je kunt het vergelijken met het gemiddelde aantal ogen van 5 dobbelstenen, al gaat het hier niet over dobbelstenen want de stochasten kunnen in deze app waarden aannemen tussen 0 en 32 (in plaats van tussen 1 en 6).

1. Wat is het verschil tussen de knop ‘5’ naast de tweede grafiek en de knop ‘N=5’ naast de derde grafiek?
2. Klik enkele keren op ‘animated’, daarna enkele keren op ‘5’, op ’10 000’ en op ‘100 000’. Beschrijf wat er gebeurt.

Experimenteer nu met de applet. Je kunt de startverdeling wijzigen, ofwel met de knop rechts ofwel door die zelf te tekenen: zet de knop op ‘custom’ en teken met de muis een verdeling. Je kunt ook het aantal stochasten waarvan je het gemiddelde neemt (‘N=...’) verhogen. Met de knop ‘fit normal’ naast de derde grafiek kun je de verdeling van het gemiddelde vergelijken met een aangepaste normale verdeling. Het is bij deze opdracht niet de bedoeling om ‘Mean’ naast de derde grafiek door iets anders te vervangen.

1. Wat stel je vast?
2. Probeer om het te doen mislopen. Kun je verdelingen maken waarbij het gemiddelde niet ‘normaal’ wordt? Zo ja, kun je verklaren waarom het misloopt?

Dit is dan een prima gelegenheid om op het ‘limiet’-aspect te wijzen. Ook al leek het bij de dobbelstenen (vorige werktekst) zo dat het histogram reeds bij een beperkt aantal dobbelstenen (stochasten) de vorm van een normale verdeling aannam, hier zie je dat 10 stochasten niet volstaan. Als je het aantal stochasten met ‘N=...’ veel groter maakt, ga je die gaten niet meer zien. De centrale limietstelling die we in de volgende deelparagraaf zullen formuleren, gaat over een limiet waarbij het aantal stochasten naar oneindig gaat.