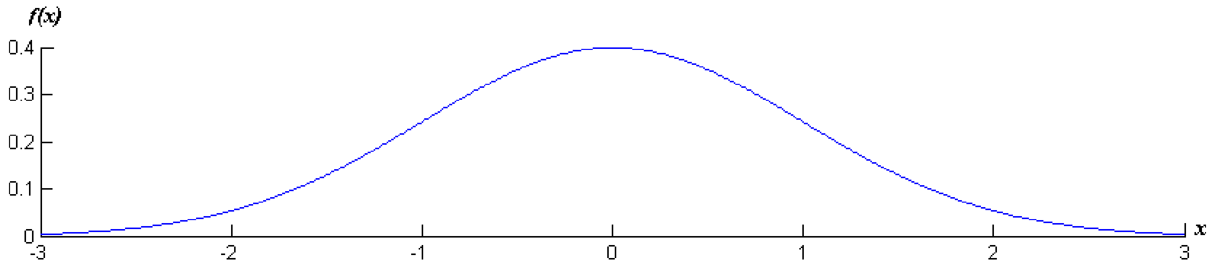
4 Chebyshev en de centrale limietstelling

4.1 Theorema van Chebyshev

Voordat men het theorema van Chebyshev kan uitleggen moet men bekend zijn met een aantal regels omtrent de relatie tussen het gemiddelde en de standaarddeviatie. Er zijn drie regels die gelden wanneer een vorm van het histogram een bel betreft:

1. Ongeveer 68% van alle waarden valt binnen 1 standaarddeviatie van het gemiddelde.
2. Ongeveer 95% van alle waarden valt binnen 2 standaarddeviaties van het gemiddelde.
3. Ongeveer 99,7% van alle waarden valt binnen 3 standaarddeviaties van het gemiddelde.



Figuur 4[a]. Belvormig histogram (bron: www.masstutor.net)

Als men het bovenstaande beschouwd kan men een meer algemene interpretatie van de standaarddeviatie toepassen, te weten het theorema van Chebyshev:

Het percentage waarden in elke steekproef dat ligt binnen *k* standaarddeviaties van het gemiddelde is tenminste

$$1-\frac{1}{k^2} \text{ voor } k>1$$

Als *k* bijvoorbeeld 3 is, dan is volgens het theorema van Chebyshev tenminste 89% ($\frac{8}{9}$) van de waarden binnen 3 standaarddeviaties van het gemiddelde liggen. Dit theorema is toepasbaar op iedere soort kansverdeling.

4.2 Theorema van de centrale limietstelling

Het theorema van de centrale limietstelling is een van de belangrijkste begrippen uit de statistiek. Deze geeft aan dat de som van een groot aantal onderling onafhankelijke en gelijk verdeelde stochastische variabelen met eindige variantie bij benadering een normale verdeling heeft[4]. Hoe groter dit aantal, hoe dichter de verdeling bij een normale verdeling licht.

[4] http://nl.wikipedia.org/wiki/Centrale_limietstelling