Was besagt der Zentrale Grenzwertsatz?

Satz (Der Zentrale Grenzwertsatz)

Wenn wir eine große Anzahl N an unabhängigen Stichprobenwerten aus einer Verteilung mit Erwartungswert μ und Varianz σ^2 ziehen, dann ist der Mittelwert der Stichprobe normalverteilt mit Erwartungswert μ und Varianz σ^2/N.

Formal: Sei $t = \frac{1}{N}\sum_{n=1}^{N} t_n$, *wobei* $t_n \overset{iid}{\sim} f$ *mit* $E(t_n) = \mu < \infty$, $Var(t_n) = \sigma^2 < \infty$.
Dann gilt:

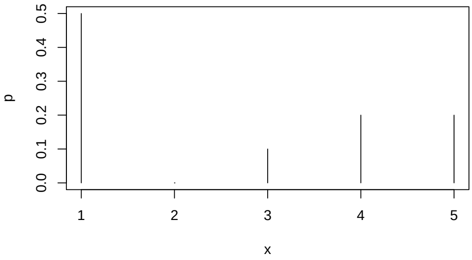
$$t \overset{a}{\sim} N(\mu, \sigma^2/N).$$

Und auch:

$$\frac{t-\mu}{\sqrt{\sigma^2/N}} = \frac{\sum_{n=1}^{N} t_n - N\mu}{\sqrt{\sigma^2 N}} \to N(0,1).$$

Die Verteilung ist beliebig?

```
x <- 1:5
p <- c(0.5,0,0.1,0.2,0.2)
plot(x, p, type = "h")
```



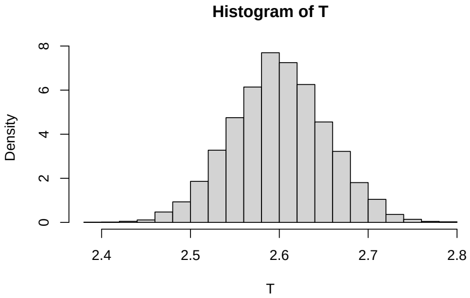

```
mu <- sum(x * p)
cat("Erwartungswert:", mu)
#> Erwartungswert: 2.6
```

```
sigma.sq <- sum((x - mu)^2 * p)
cat("Varianz:", sigma.sq)
#> Varianz: 2.84
```

Und wir kommen wieder bei der Normalverteilung an?

```r
N <- 1000
t <- function()
  mean(
    sample(
      x = x, size = N,
      replace = TRUE, prob = p
    )
  )
replicate(3, t())
#> [1] 2.506 2.521 2.535
```

```r
T <- replicate(10000, t())
hist(T, breaks = 20, freq = FALSE)
```

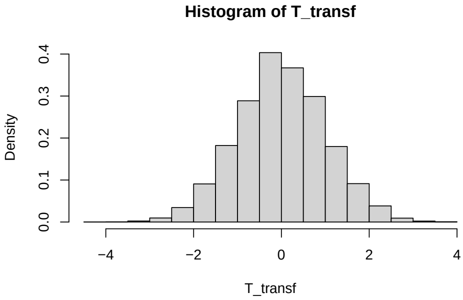


Und auch bei der Standardnormalverteilung?

> Satz (Der Zentrale Grenzwertsatz)
>
> *[...]*
>
> *Dann gilt:* $t \overset{a}{\sim} N(\mu, \sigma^2/N)$.
>
> *Und auch:* $\frac{t-\mu}{\sqrt{\sigma^2/N}} \to N(0,1)$.

```
T_transf <- (T - mu) /
  sqrt(sigma.sq / N)
```

```
hist(T_transf, breaks = 20,
     freq = FALSE)
```



Und wofür kann man das verwenden?

[...] unabhängige Stichprobe $X_1, \ldots, X_N$ [...] Erwartungswert jeder Ziehung beträgt μ und Varianz σ^2 [...] Wie hoch ist die Wahrscheinlichkeit, dass das Stichprobenmittel höchstens x beträgt?

$$P\left(\frac{1}{N}\sum_{i=1}^{N} X_i \leq x\right) = P\left(\frac{\frac{1}{N}\sum_{i=1}^{N} X_i - \mu}{\sqrt{\sigma^2/N}} \leq \frac{x-\mu}{\sqrt{\sigma^2/N}}\right)$$

$$\approx \Phi\left(\frac{x-\mu}{\sqrt{\sigma^2/N}}\right)$$