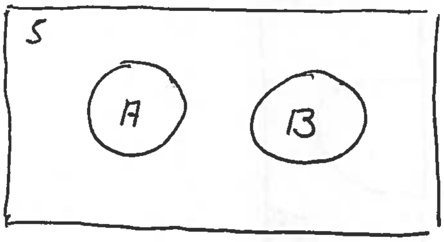
ST 1101

SANNSYNSREKNING	FORDELINGAR	INFERENS
SANNSYN	TILFELDIGE VARIABLE	SENTRALGRENSE-TEOREMET
ADDISJONSSETNINGA	FORVENTNING	ESTIMERING
MULTIPLIKASJONS-SETNINGA	VARIANS	KONFIDENSINTERVALL
BETINGA SANNSYN	STANDARDAVVIK	HYPOTESETESTING
TOTAL SANNSYN	KOVARIANS	
BAYES FORMEL	UNIFORM	
KOMBINATORIKK	BINOMISK	
	GEOMETRISK	
	NEGATIVT BINOMISK	
	HYPERGEOMETRISK	
	POISSON	
	NORMAL	
	EKSPONENTIAL	
	GAMMA	

SANNSYN

STOKASTISKE FORSØK

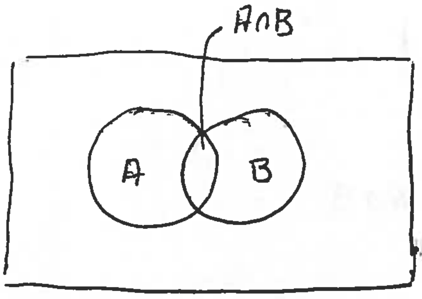


$$P(S) = 1$$

$$P(A) \geq 0$$

$$P(A \cup B) = P(A) + P(B), \quad A \cap B = \emptyset$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \quad A_i \cap A_j = \emptyset, \ \forall_{ij} \ i \neq j$$



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

BETINGA SANNSYN:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

UAVHENGIGHEIT:

$$P(A|B) = P(A), \quad P(A|B^c) = P(A)$$

$$P(A \cap B) = P(A) \cdot P(B)$$

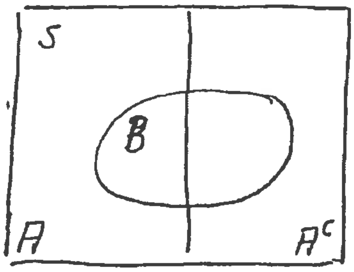
$$P(A_1 \cap \cdots \cap A_m) = \prod_{i=1}^{m} P(A_i)$$

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_j})$$

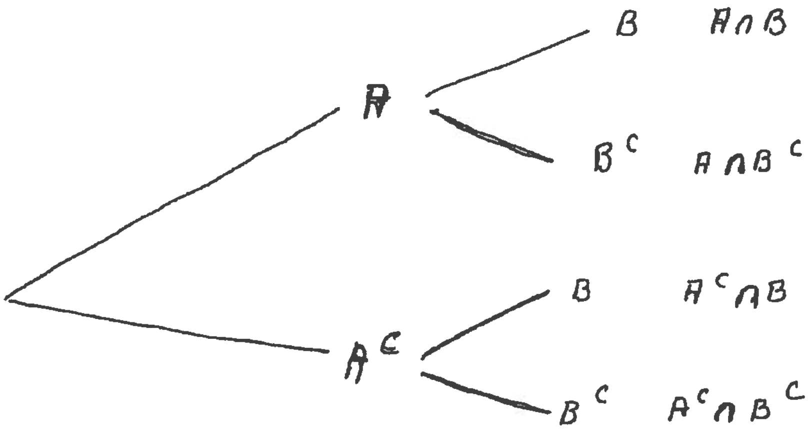
$\forall_j$ og $i_1, \ldots, i_j$

BAYES REGEL

TOTAL SANNSYN



$$P(B) = P(A \cap B) + P(A^c \cap B)$$



$$P(A|B) = \frac{P(A \cap B)}{P(A \cap B) + P(A^c \cap B)}$$ Bayes regel

$$= \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c}$$

KOMBINATORIKK

UNIFORME SANNSYNSMODELLAR $P(A) = \frac{g}{m}$

MULTIPLIKASJONSREGELEN: $m_1 \cdot m_2 \cdot \dots \cdot m_k$.

TALET PÅ ORDNA UTVAL $\frac{m!}{(m-s)!} = m(m-1)\dots(m-s+1)$

TALET PÅ IKKJE ORDNA UTVAL $\frac{m!}{(m-s)!\,s!} = \binom{m}{s}$

FORDELINGSFUNKSJONAR

Diskret

$$F_X(x) = P(X \le x) = \sum_{u \le x} P(X = u)$$

$X = k$ heiltal

$$F_X(k) = P(X \le k) = \sum_{i \le k} P(X = i)$$

$$P(X > k) = P(X \ge k+1) = \sum_{i=k+1}^{\infty} P(X = i)$$

$$P(X = k) = P(X \le k) - P(X \le k-1)$$

Kontinuerlig

$$F_Y(y) = P(Y \le y) = \int_{-\infty}^{y} f_Y(y)\,dy$$

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$

$$P(Y = y) = 0$$

FLEIRDIMENSJONALE FORDELINGSFUNKSJONAR

Diskret

$$F_{X,Y}(x,y) = P(X \le x \cap Y \le y) = \sum_{v \le y} \sum_{u \le x} P(X=u \cap Y=v)$$

$$P(X=x) = \sum_{\text{alle } y} P(X=x \cap Y=y)$$

Kontinuerleg

$$F_{X,Y}(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v)\, du\, dv$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$

Uavh.

$$P(X_1 = x_1 \cap X_2 = x_2 \cap \cdots \cap X_m = x_m) = \prod_{i=1}^{m} P(X_i = x_i)$$

$$f_{Y_1, \ldots, Y_m}(y_1, \ldots, y_m) = \prod_{i=1}^{m} f_{Y_i}(y_i)$$

FORVENTING TIL EIN FUNKSJON

$$E[g(X)] = \begin{cases} \sum g(x_i)\, P(X=x_i), & X \text{ diskret} \\ \int_{-\infty}^{\infty} g(x) f(x)\, dx, & X \text{ kontinuerleg} \end{cases}$$

$$E[g(X,Y)] = \begin{cases} \sum_y \sum_x g(x,y)\, P(X=x \cap Y=y), & \text{Diskret} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\, dx\, dy, & \text{kontinuerleg} \end{cases}$$

FORVENTING OG VARIANS

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} x\, f(x)\, dx, & X \text{ kontinuerleg} \\ \sum_i x_i\, P(X = x_i), & X \text{ diskret} \end{cases}$$

$$Var[X] = \begin{cases} \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx, & X \text{ kontinuerleg} \\ \sum_i (x_i-\mu)^2\, P(X = x_i), & X \text{ diskret} \end{cases}$$

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X\, \sigma_Y}$$

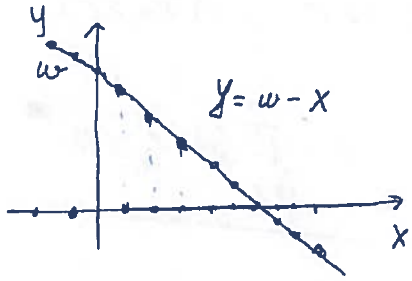
$$E\left[\sum a_i X_i\right] = \sum a_i E[X_i]$$

$$Var\left[\sum a_i X_i\right] = \sum a_i^2 Var[X_i] + 2 \sum_{i<j}\sum a_i a_j Cov(X_i, X_j)$$

KONVOLUSJONSSETNINGA

$W = X + Y = g(X,Y)$ $\qquad$ X, Y uavh.

$$P(W=w) = \sum_{\text{alle } x} P(X=x \cap Y=w-x)$$



Diskret

Kontinuerleg

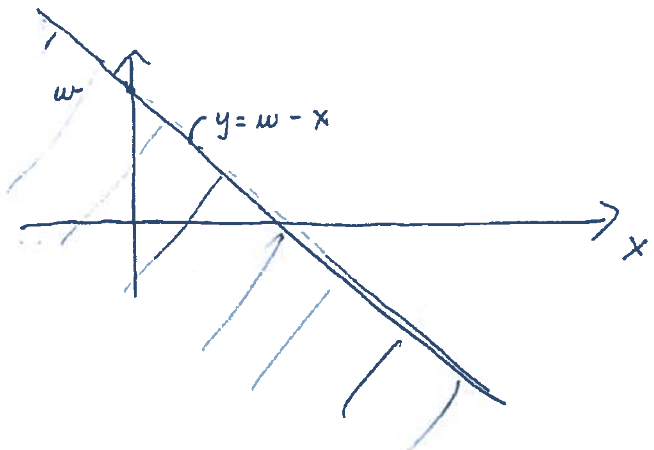
X, Y uavh.

$W = X + Y$

$$F_W(w) = P(W \le w) = P(X+Y \le w) = P(Y \le w - X)$$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{w-x} f_X(x) f_Y(y)\, dy \right) dx = \int_{-\infty}^{\infty} f_X(x) F_Y(w-x)\, dx$$

$$\frac{d F_W(w)}{dw} = \int_{-\infty}^{\infty} f_X(x) f_Y(w-x)\, dx$$



MAX OG MIN

$X_1, \dots X_n$ uavhengige

$$P(\max(X_1, \dots X_n) \leq x) = P(X_1 \leq x \cap X_2 \leq x \cap \dots \cap X_n \leq x)$$

$$= \prod_{i=1}^{n} F_{X_i}(x)$$

$$P(\min(X_1, \dots, X_n) \leq x) = 1 - P(X_1 > x \cap X_2 > x \cap \dots \cap X_n > x)$$

$$= 1 - \prod_{i=1}^{n} P(X_i > x) = \underline{1 - \prod_{i=1}^{n}(1 - F_{X_i}(x))}$$

BETINGA FORDELINGAR

$$P(Y=y \mid X=x) = \frac{P(Y=y \cap X=x)}{P(X=x)}, \quad \text{Diskret}$$

$$f_{Y|X=x} = \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad \text{kontinuerleg}$$

SENTRALGRENSETEOREMET

$X_1, \dots, X_n$ tilf. utvalg $\begin{cases} \text{uavh.} \\ \text{identisk fordelte} \end{cases}$

$E[X_i] = \mu, \qquad Var[X_i] = \sigma^2$

$$E[\bar{X}] = \mu$$

$$Var[\bar{X}] = \frac{\sigma^2}{n}$$

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \to \infty]{} N(0,1)$$

eller: $\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$

eller: $\sum_{i=1}^{n} X_i \approx N(n\mu, n\sigma^2)$

DISKRETE FORDELINGAR

Binomisk (n, p)

X = talet på gonger A skjer i n forsøk

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$E[X] = np, \quad Var[X] = np(1-p)$$

Geometrisk (p)

X = talet på forsøk til A skjer 1. gong

$$P(X=k) = (1-p)^{k-1} p$$

$$E[X] = \frac{1}{p}, \quad Var[X] = \frac{1-p}{p^2}$$

Negativ Binomisk (r, p)

X = talet på forsøk til A skjer r-te gong.

$$P(X=k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

$$E[X] = \frac{r}{p}, \quad Var[X] = \frac{r(1-p)}{p^2}$$

Hypergeometrisk

X = talet av type A i eit utval av storleik n frå N der r stykker er av type A

$$P(X=k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$$

$$E[X] = n\frac{r}{N}, \quad Var[X] = \frac{N-n}{N-1}\, \frac{nr}{N}\left(1-\frac{r}{N}\right)$$

TILNÆRMINGAR/RELASJONAR

p liten, $np = \lambda$

$$P(X=k) = \frac{(np)^k e^{-np}}{k!}$$

$$\left.\begin{matrix} np \geq 5 \\ n(1-p) \geq 5 \end{matrix}\right\} \Rightarrow X \approx N(np, np(1-p))$$

X_i, $i = 1, 2, \ldots, r$ uavh. og geometrisk (p)

$$\sum_{i=1}^{r} X_i \sim \text{Negativ binomisk } (r, p)$$

$$\approx N\left(\frac{r}{p}, \frac{r(1-p)}{p^2}\right)$$

$N \gg n \Rightarrow$

$$P(X=k) \approx \binom{n}{k}\left(\frac{r}{N}\right)^k \left(1-\frac{r}{N}\right)^{n-k}$$

$\approx$

Poisson (λt)

X = talet på gonger A skjer over tid (på flate, i volum o.l.)

$$P(X=k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

$E[X] = \lambda t$, $Var[X] = \lambda t$

KONTINUERLEGE FORDELINGAR

Normal (μ, σ^2)

$$f(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

$-\infty < y < \infty$

$E[Y] = \mu$, $Var[Y] = \sigma^2$

Eksponential (λ)

$$f(y;\lambda) = \begin{cases} \lambda e^{-\lambda y}, & y > 0 \\ 0, & \text{elles.} \end{cases}$$

$E[Y] = \frac{1}{\lambda}$, $Var[Y] = \frac{1}{\lambda^2}$

Gammafordeling (λ, n)

$$f(y,\lambda,n) = \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y}, \quad y > 0$$

$E[Y] = \frac{n}{\lambda}$, $Var[Y] = \frac{n}{\lambda^2}$

TILNÆRMINGAR / RELASJONAR

$\lambda t \geq 18$

$X \approx N(\lambda t, \lambda t)$

Ventetid til 1. hending (mellom 2 hendingar) i Poisson-prosessen.

n heiltal: Ventetid til n-te hending (mellom ei hending og den n-te neste) i Poissonprosessen.

$Y_1, Y_2, \ldots, Y_n$ eksponentialfordelte (λ) og uavh. $\Rightarrow \sum_{i=1}^{n} Y_i \sim \text{gamma}(\lambda, n)$

$\approx N\left(\frac{n}{\lambda}, \frac{n}{\lambda^2}\right)$

MOMENT GENERERANDE FUNKSJONAR

$$M_X(t) = E[e^{tX}]$$

$$M_{aX+b}(t) = e^{bt} M_X(at)$$

TO RESULTAT

$$E[X^n] = \frac{d^n}{dt^n} M_X(t) \Big|_{t=0}$$

$$M_{\sum_{i=1}^{m} X_i}(t) = \prod_{i=1}^{m} M_{X_i}(t) \quad \text{uavh.}$$

$X_1, X_2, \ldots, X_m$ uavh.

$X_i \sim \text{Poisson}(\lambda_i) \Rightarrow \sum_{i=1}^{m} X_i \sim \text{Poisson}\left(\sum_{i=1}^{m} \lambda_i\right)$

$X_i \sim B(n_i, p) \Rightarrow \sum_{i=1}^{m} X_i \sim B\left(\sum_{i=1}^{m} n_i, p\right)$

$X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow \sum_{i=1}^{m} X_i \sim N\left(\sum_{i=1}^{m} \mu_i, \sum_{i=1}^{m} \sigma_i^2\right)$

$X_i \sim \text{geometrisk}(p) \Rightarrow \sum_{i=1}^{m} X_i \sim \text{Negativ binomisk}(m, p)$

$X_i \sim \text{negativ binomisk}(n_i, p) \Rightarrow \sum_{i=1}^{m} X_i \sim \text{Negativ binomisk}\left(\sum_{i=1}^{m} n_i, p\right)$

$X_i \sim \text{eksponentiell fordelt}(\lambda) \Rightarrow \sum_{i=1}^{m} X_i \sim \text{gammafordelt}(m, \lambda)$

$X_i \sim \text{gammafordelt}(n_i, \lambda) \Rightarrow \sum_{i=1}^{m} X_i \sim \text{gammafordelt}\left(\sum_{i=1}^{m} n_i, \lambda\right)$

INFERENS

$X_1, \ldots, X_n$ u.o.f.

$$L(\theta) = \begin{cases} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\theta) = \prod_{i=1}^{n} f_{X_i}(x_i;\theta) \\ P(X_1=x_1, X_2=x_2, \ldots, X_n=x_n; \theta) = \prod_{i=1}^{n} P(X_i=x_i;\theta) \end{cases}$$

SSME: θ_e maksimerer $L(\theta)$, evt. $\ln L(\theta)$

$\theta_e = \theta(x_1, \ldots, x_n)$ estimat.

$\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ estimator.

Momentestimatorane: $E[Y^k] = \frac{1}{n}\sum_{i=1}^{n} y_i^k, \quad k=1,2,\ldots$

Konfidensintervall

$$P\left(-z_{\frac{\alpha}{2}} \le \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\bar{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \quad \bar{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Hypotesetesting i normalfordeling

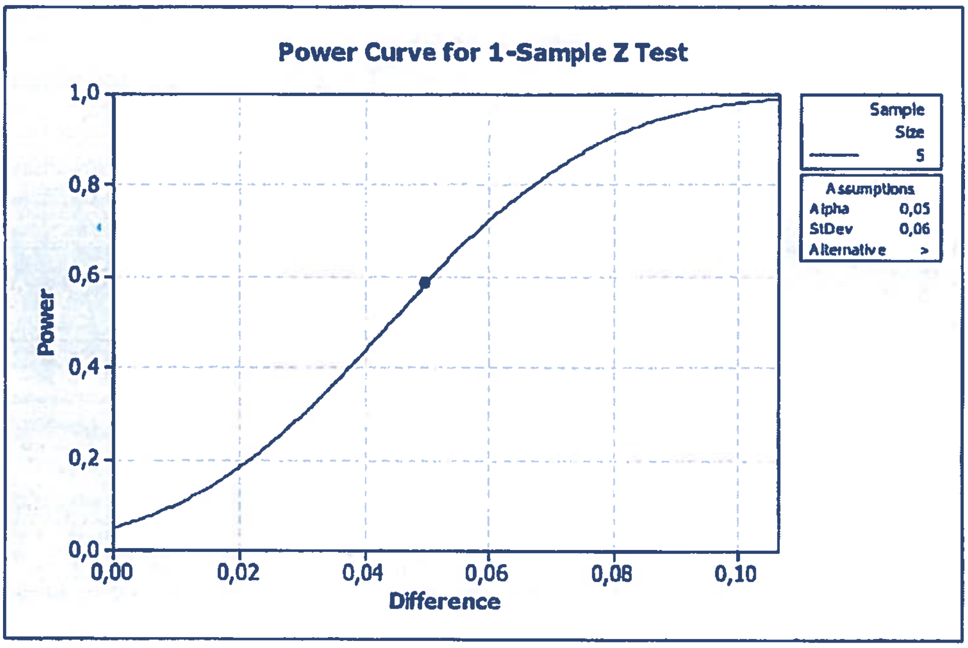
$$H_0: \mu \begin{Bmatrix} \leq \\ = \\ \geq \end{Bmatrix} \mu_0 \quad H_1: \mu \begin{Bmatrix} > \\ \neq \\ < \end{Bmatrix} \mu_0$$

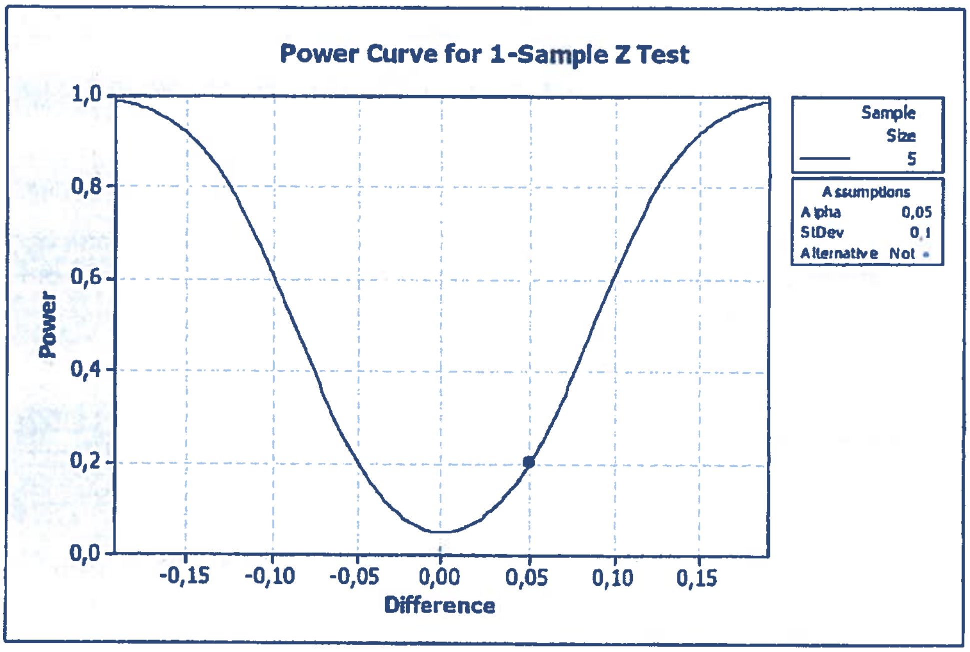
$$\text{Forkast } H_0 \text{ om } \begin{Bmatrix} Z = \dfrac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \geq z_\alpha \\ |Z| = \left| \dfrac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \right| \geq z_{\frac{\alpha}{2}} \\ Z = \dfrac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq -z_\alpha \end{Bmatrix}$$

P-verdien er sannsynet for å få noko som er minst like ekstremt som det vi har vi har observert gitt H_0.

P-verdien er det minste signifikansnivået som fører til forkastning av H_0.

$$P(Z \geq z_{obs} | H_0), \;\; 2P(Z \geq |z_{obs}| | H_0), \;\; P(Z \leq z_{obs} | H_0)$$


Power Curve for 1-Sample Z Test
1,0
0,8
0,6
0,4
0,2
0,0
Power
0,00
0,02
0,04
0,06
0,08
0,10
Difference
Sample Size
5
Assumptions
Alpha 0,05
StDev 0,06
Alternative >


Power Curve for 1-Sample Z Test
1,0
0,8
0,6
0,4
0,2
0,0
Power
-0,15
-0,10
-0,05
0,00
0,05
0,10
0,15
Difference
Sample Size
5
Assumptions
Alpha 0,05
StDev 0,1
Alternative Not =