Name: ………………………………….

Matrikel-Nr.: ……………………………

Hinweise:

- Diese Klausur enthält 1 Aufgabe.
- Die Unterlagen bestehen aus 12 Seiten inkl. Deckblatt.
- Die Dauer der Klausur beträgt 180 Minuten.
- Erlaubte Hilfsmittel:
  - Vorlesungsskript, Vorlesungsfolien
  - Schriftliche Unterlagen
  - Zugewiesener Laborrechner
  - Taschenrechner
- Notieren Sie auf jedem Blatt Ihren Namen und Matrikelnummer.
- Stellen Sie Ihren Lösungsweg ausführlich dar und begründen Sie Ihr Vorgehen.
- Geben Sie dieses Deckblatt, die Aufgabenstellung und alle bearbeiteten Blätter mit Ihrer Klausur ab.

max. Punktzahl	Ist-Punktzahl
100	

Note	

Beschreibung der Aufgabe

In einem Tagungsbeitrag am Deutschen Zentrum für Luft- und Raumfahrt (DLR) geht Dr. Uwe Waschatz (Volkswagen AG) auf den Einsatz der statistischen Versuchsplanung in der Motor-applikation ein. Dabei untersucht er unter anderem, wie sich die Relative Luftmasse R und die Drehzahl n auf das Schleppmoment des Motors M auswirken.

Quelle: www.dlr.de/fs/ResourceImage.aspx?raid=20558

Die Eingangsgrößen und die Ausgangsgröße sind in den folgenden Tabellen zusammengefasst.

Eingangsgröße	**Variable**	**Einheit**
Relative Luftmasse	R	%
Drehzahl	n	U/min

Ausgangsgröße	**Variable**	**Einheit**
Schleppmoment	M	N·m

In den folgenden Aufgaben werden seine Untersuchungen nachvollzogen.

Signifikanzbewertung von Eingangsgrößen

Um den Versuchsraum einzugrenzen, werden Voruntersuchungen durchgeführt. Es wird die Hypothese aufgestellt, dass die Drehzahl n keinen signifikanten Einfluss auf das Schleppmoment M besitzt. Deshalb wird bei einer Relativen Luftmasse von R = 20 % ein Versuch durchgeführt, bei dem das Schleppmoment M bei unterschiedlichen Drehzahlen n bewertet wird. Die Daten sind im Datensatz *Signifikanz.mat* gespeichert.

Schleppmoment M / N·m			
n_1 = 1000 U/min		n_2 = 2140 U/min	
- 12.9564	- 13.0070	- 13.0719	- 13.0348
- 13.0567	- 12.9638	- 12.9638	- 12.9203
- 13.0136	- 12.9495	- 13.0251	- 12.9518
- 12.9634	- 13.1092	- 12.8958	- 13.0476
- 12.8836	- 13.1695	- 13.0652	- 13.0464

a) Ist die Drehzahl n signifikant für das Schleppmoment? Führen Sie mit den angegebenen Daten einen Hypothesentest auf gleichen Mittelwert durch. Wie lautet die Nullhypothese, wie die Alternative? Welche Zufallsvariable wählen Sie für den Hypothesentest?

b) Erstellen Sie für den Hypothesentest die zugehörige Gütefunktion. Welche Abweichung können Sie mit einer Wahrscheinlichkeit von 99 % erkennen?

c) Plausibilisieren Sie das Ergebnis des Hypothesentests mit einer geeigneten grafischen Darstellung.

d) Können Sie mit Hilfe einer Korrelationsanalyse bewerten, ob die Drehzahl n einen signifikanten Einfluss aufweist? Führen Sie auch dazu einen geeigneten Hypothesentest durch. Wie lautet die verwendete Zufallsvariable?

Statistische Versuchsplanung und Regressionsrechnung

Über einen Versuch soll der Zusammenhang zwischen den Eingangsgrößen Relative Luftmasse R und Drehzahl n und dem Schleppmoment M ermittelt werden. Die Eingangsgrößen und ihre Einstellbereiche sind in der folgenden Tabelle zusammengefasst.

Eingangsgröße	Variable	Einheit	Minimum	Maximum
Relative Luftmasse	R	%	20	80
Drehzahl	n	U/min	800	5000

Zur Ermittlung der Versuchsergebnisse wird die MATLAB-Funktion *Schleppmoment.p* verwendet. Sie ermittelt das Schleppmoment M an der Stelle X, die sich aus den Variablen X = [R n] zusammensetzt. Alle Variablen können auch als Spaltenvektor eingesetzt werden. Der Aufruf sieht beispielsweise so aus: Schleppmoment([20 2000]).

e) Stellen Sie einen vollfaktoriellen 3^K - Versuchsplan auf.

f) Prüfen Sie die Einstellungen der Eingangsgrößen in normierter Form. Ist der Versuchsplan orthogonal? Führen Sie dieselbe Prüfung mit den physikalischen Einstellwerten durch. Diskutieren Sie die Ergebnisse im Hinblick auf die angestrebte Orthogonalität.

g) Bestimmen Sie die Versuchsergebnisse und speichern Sie das Versuchsergebnis ab. Verwenden Sie für die folgenden Aufgabenteile diesen Datensatz weiter.

h) Stellen Sie eine vollquadratische Regressionsfunktion auf und reduzieren Sie die Regressionsfunktion auf die signifikanten Terme ohne Rücksicht auf das adjungierte Bestimmtheitsmaß. Wie ändert sich das adjungierte Bestimmtheitsmaß?

Zu Validierung des Ergebnisses werden Versuche mit den Einstellungen durchgeführt, die in der folgenden Tabelle zusammengefasst sind. Die Daten sind im Datensatz *Reststreuungsanalyse.mat* gespeichert.

Größe	1	2	3	4	5	6	7	8	9	10
Relative Luftmasse R / %	43.7	65.7	58.4	69.5	45.9	69.9	42.6	28.5	54.1	49.4
Drehzahl n / U/min	3073	2452	2566	1376	2027	831	4717	3359	4546	2567
Schleppmoment M / N·m	-14.1	-11.5	-12.2	-9.2	-11.4	-9.8	-17.7	-15.2	-17.1	-12.4

i) Führen Sie für diesen Datensatz eine Reststreuungsanalyse durch. Erkennen Sie einen funktionalen Zusammenhang zwischen Residuen und Einstellgrößen? Wie würden Sie Ihr Modell ändern, um den Betrag der Residuen zu verringern?

j) Stellen Sie einen D-optimalen Versuchsplan für dieses neue Modell auf und bestimmen Sie für dieses Modell die Regressionsparameter. Wie wirkt sich die Änderung auf die Reststreuungen aus?

Statistische Tolerierung

Zur Bestimmung der Relativen Luftmassen wird der Luftmassenstrom dm/dt über ein sogenanntes Saugrohrmodell bestimmt, das auf der allgemeinen Gasgleichung basiert. Durch Ableiten nach der Zeit ergibt sich für den Luftmassenstrom der Zusammenhang

$$\frac{dm}{dt} = \frac{p}{R \cdot T} \cdot \eta \cdot V_H \cdot n$$

In die Berechnung des Luftmassenstroms gehen damit die Größen ein, die in der folgenden Tabelle zusammengefasst sind.

Variable	Einheit	Arbeits-punkt	Toleranz		Bemerkungen
			Minimum	Maximum	
Saugrohrdruck p	Pa	101300	- 0.5 %	+ 0.5 %	Normalverteilung, Toleranz entspricht $\pm 3 \cdot \sigma$
Gaskonstante R	J/(kg·K)	287.1			Konstante, keine Toleranz
Temperatur T	K	323	- 1 K	+ 1 K	Normalverteilung, Toleranz entspricht $\pm 3 \cdot \sigma$
Füllungsgrad η	%	80	- 2	+ 2	Gleich-verteilung
Hubvolumen V_H	m³	$1.9 \cdot 10^{-3}$	$- 5 \cdot 10^{-6}$	$5 \cdot 10^{-6}$	Gleich-verteilung
Drehzahl n	U/h	$1.5 \cdot 10^{5}$	- 0.3 %	+ 0.3 %	Gleich-verteilung

Alle Größen sind statistisch voneinander unabhängig und weisen untereinander keine signifikante Korrelation auf.

k) Führen Sie eine statistische Tolerierung auf Basis des Grenzwertsatzes der Wahrscheinlichkeitsrechnung durch. Wie groß ist die Toleranz des Luftmassenstroms bei einer Konfidenzzahl von γ = 99.73 %?

l) Verifizieren Sie das Ergebnis durch eine statistische Simulation. Wie groß ist die Toleranz des Luftmassenstroms bei einer Konfidenzzahl von γ = 99.73 %?

m) Vergleich Sie die Ergebnisse der beiden Tolerierungsmethoden.

n) Wodurch wird die Toleranz des Luftmassenstroms maßgeblich bestimmt? Ordnen Sie die Eingangsgrößen

Musterlösung

a) Der t-Test zum Vergleich zweier normalverteilter Stichproben mit identischem Stichprobenumfang N auf gleichen Mittelwert wird mit der Zufallsvariable

$$t = \frac{\bar{x}_2 - \bar{x}_1 - (\mu_2 - \mu_1)}{\sqrt{\frac{1}{N} + \frac{1}{N}} \cdot s} = \frac{\Delta x - \Delta\mu}{\sqrt{\frac{2}{N}} \cdot s}$$

durchgeführt. Die Nullhypothese H0 lautet $\Delta\mu = 0$, damit ergibt sich

$$t = \frac{\bar{x}_2 - \bar{x}_2}{\sqrt{\frac{2}{N}} \cdot s} = \frac{\Delta x}{\sqrt{\frac{2}{N}} \cdot s}$$

mit

$$s^2 = \frac{s_1^2 \cdot (N-1) + s_2^2 \cdot (N-1)}{N + N - 2} = \frac{s_1^2 + s_2^2}{2}$$

Zur Prüfung der Nullhypothese wird der Annahmebereich berechnet. Zur Berechnung der Konstanten

$$c_1 = F^{-1}\left(\frac{1-\gamma}{2}\right)$$

und

$$c_2 = F^{-1}\left(\frac{1+\gamma}{2}\right)$$

wird die t-Verteilung mit 2·N - 2 Freiheitsgraden verwendet. Damit ergibt sich der Annahmebereich zu

$$\gamma = P\left(c_1 \cdot \sqrt{\frac{2}{N}} \cdot s < \bar{x}_2 - \bar{x}_2 \le c_2 \cdot \sqrt{\frac{2}{N}} \cdot s\right) = P\left(-2.1009 \cdot \sqrt{\frac{2}{10}} \cdot 0.0748 < \bar{x}_2 - \bar{x}_2 \le 2.1009 \cdot \sqrt{\frac{2}{10}} \cdot 0.0748\right)$$

$$= P\left(-0.0703 < \bar{x}_2 - \bar{x}_2 \le 0.0703\right) = P\left(x_{C1} < \bar{x}_2 - \bar{x}_2 \le x_{C2}\right)$$

Die Differenz der Mittelwerte bei der vorliegenden Stichprobe beträgt 0.0050 N·m, sie liegt im Annahmebereich. Die Nullhypothese auf gleiche Mittelwerte wird deshalb angenommen.

```
% Laden der Daten
load Signifikanz.mat

% Bestimmung des Annahmebereichs des Hypothesentests auf gleiche Mittelwerte
N = 10;
gamma = 0.95
c1 = tinv((1-gamma)/2,2*N-2)
c2 = tinv((1+gamma)/2,2*N-2)
s = sqrt((var(M1)+var(M2))/2)
dxmin = c1*sqrt(2/N)*s
dxmax = c2*sqrt(2/N)*s
dx = mean(M2) - mean(M1)
```

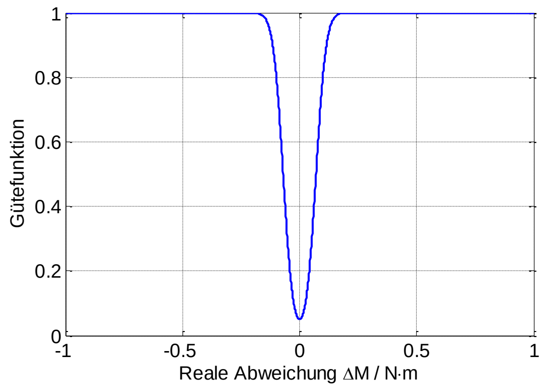
b) Mit Aufgabe a) liegen die Annahmegrenzen x_{C1} und x_{C2} vor. Die Bestimmung der Gütefunktion muss auf die t-Verteilung mit 2·N - 2 Freiheitsgraden zurückgeführt werden. Mit der t-verteilten Variable

$$t = \frac{\Delta x - \Delta\mu}{\sqrt{\frac{2}{N}} \cdot s}$$

ergibt sich bei variablem $\Delta\mu$ die Gütefunktion

$$Q = F\left(\frac{x_{C1} + \Delta\mu}{\sqrt{\frac{2}{N}} \cdot s}\right) + 1 - F\left(\frac{x_{C2} + \Delta\mu}{\sqrt{\frac{2}{N}} \cdot s}\right)$$

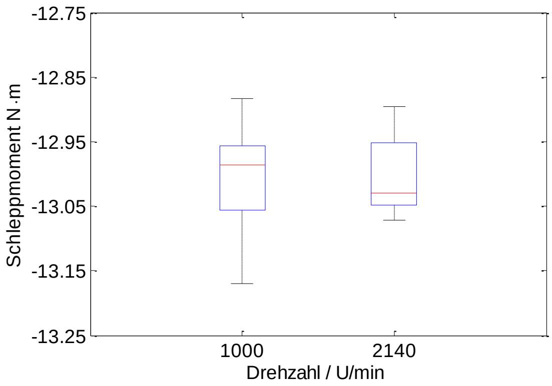
Sie ist in dem folgenden Diagramm dargestellt.



Eine Abweichung von $\Delta M = \pm$ 0.1550 N·m kann mit einer Sicherheit von 99 % erkannt werden.

```
%% Aufstellen und Darstellen der Gütefunktion
dmu = -1:0.001:1;
G = tcdf((dxmin+dmu)/sqrt(2/N)/s,2*N-2) + 1 - ...
    tcdf((dxmax+dmu)/sqrt(2/N)/s,2*N-2);

% Grafische Darstellung
plot(dmu,G,'b','Linewidth',2);
axis([-1 1 0 1]);
set(gca,'XTick',-1:0.5:1,'YTick',0:0.2:1);
xlabel('Reale Abweichung \DeltaM / N\cdotm');
ylabel('Gütefunktion');

% Numerische Auswertung
bin = find(G<=0.99,1,'first');
dM = dmu(bin)
```

c) Zur Plausibilisierung wird ein Boxplot verwendet.



Der Boxplot zeigt eine Überlappung der Inter-Quartil-Abstände beider Stichproben, die Signifikanzbewertung wird damit bestätigt.

```
% Plausibilisierung mit Boxplot
boxplot([M1 M2],'labels',{'1000' '2140'});
axis([0 3 -13.25 -12.75]);
set(gca,'YTick',-13.25:0.1:-12.75);
xlabel('Drehzahl / U/min ');
ylabel('Schleppmoment N \cdotm');
```

d) Die Korrelation bewertet, ob steigende Werte der Eingangsgröße typischerweise mit steigenden Werten der Ausgangsgröße auftreten. Sie kann deshalb zur Signifikanzbewertung herangezogen werden. Es wird geprüft, ob der Korrelationskoeffizient signifikant von null abweicht. Dazu wird die Zufallsvariable

$$t = r \cdot \sqrt{\frac{N-2}{1-r^2}}$$

verwendet. Die Auswertung erfolgt mit dem MATLAB-Befehl corrcoef. Es liegt eine schwache Korrelation von r = 0.0352 vor, der p-Value liegt mit 88.3 % deutlich über der Grenze von 5 %, die Korrelation ist nicht signifikant. Damit unterscheiden sich die beiden Stichproben nicht signifikant.

```
% Signifikanzbewertung mit der Korrelation
[R,P,RLO,RUP] = corrcoef([[n1;n2] [M1;M2]])
```

e) Der vollfaktorielle 3^K-Versuchpan wird in MATLAB erzeugt und in physikalische Daten gewandelt.

```
% Vollfaktorielle 3K-Versuchsplan
ff_design = fullfact([3 3])-2;
Xff = [ff_design(:,1)*30+50, ff_design(:,2)*2100+2900];
```

f) Zur Prüfung der Orthogonalität der Einstellgrößen wird für die normierten Einstellungen und die physikalischen Einstellungen die Kovarianzmatrix aufgestellt. Es zeigt sich, dass die Einstellgrößen in der normierten Form orthogonal sind, während die physikalischen Größen nicht orthogonal sind. Ursache ist eine nichtlineare Abbildung zwischen beiden Größe über das Minimum und Maximum jeder einzelnen Variable.

```
% Prüfung auf Orthogonalität
Test_Normiert = ff_design'*ff_design
Test_Physikalsch = Xff'*Xff
```

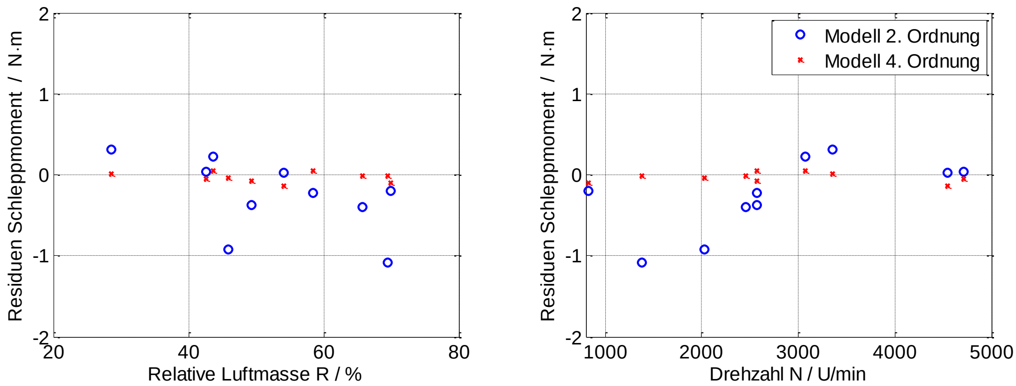
g) Berechnung der Ausgangsgröße und Abspeichern des Datensatzes.

```
% Berechnung der Versuchsergebnisse und Regression
% M_ff = Schleppmoment(X_ff);
% save VollfaktoriellerVersuch.mat M_ff X_ff;
load VollfaktoriellerVersuch.mat;
```

h) Für das Aufstellen der Regressionsfunktion wird die Funktion regstats verwendet. Es zeigt sich, dass der quadratische Term für die Relative Luftmasse nicht signifikant ist, alle anderen Terme sind signifikant. Durch das Entfernen des Terms steigt das adjungierte Bestimmtheitsmaß leicht an.

```
%% Regressionsfunktion und Reduktion um nichtsignifikante Terme
model = [0 0; 1 0; 0 1; 1 1; 2 0; 0 2];
stats = regstats(M_ff,X_ff,model);
stats.tstat.pval
stats.adjrsquare

model = [0 0; 1 0; 0 1; 1 1; 0 2];
stats = regstats(M_ff,X_ff,model);
stats.tstat.pval
stats.adjrsquare
b_ff = stats.tstat.beta
```

i) Für die gegebenen Einstellungen wird die Regression berechnet. Die Differenz zwischen den berechneten und den gemessenen Werten sind die Residuen, sie werden als Funktion der beiden Eingangsgrößen dargestellt.



Es zeigt sich, dass die Residuen systematisch von der Drehzahl abhängig sind, während bei der Relativen Luftmasse kein systematischer Zusammenhang besteht. Deshalb wird die Regressionsfunktion um Terme dritter und vierte Ordnung erweitert.

```
% Reststreungsanalyse erstellen und darstellen
Mreg = b_ff(1) + b_ff(2)*Rval + b_ff(3)*Nval + b_ff(4)*Rval.*Nval +
b_ff(5)*Nval.^2;
Mres = Mreg - Mval;

% Grafische Darstellung
subplot(1,2,1);
plot(Rval,Mres,'bo','linewidth',2);
axis([20 80 -2 2]);
set(gca,'XTick',20:20:80,'YTick',-2:1:2);
xlabel('Relative Luftmasse R / %');
ylabel('Residuen Schleppmoment  /  N \cdotm');

subplot(1,2,2);
plot(Nval,Mres,'bo','linewidth',2);
axis([800 5000 -2 2]);
set(gca,'XTick',0:1000:5000,'YTick',-2:1:2);
xlabel('Drehzahl N / U/min');
ylabel('Residuen Schleppmoment  /  N \cdotm');
```

j) Für das neue Modell wird ein D-optimaler Versuchsplan erstellt und berechnet. Eine Prüfung der Regressionsterme zeigt, dass alle Terme signifikant sind. Das adjungierte Bestimmtheitsmaß ist auf 99.96 % gestiegen. Die Residuen nach der Verbesserung sind in das Diagramm oben bereits eingetragen, sie sind kleiner als bei dem Modell zweiter Ordnung und weisen keinen systematischen Zusammenhang mehr auf.

```
% D-optimaler Versuchsplan
model = [0 0; 1 0; 0 1; 1 1; 0 2; 0 3; 0 4];
AnzahlVariable = 2;
AnzahlVersuche = 9;
dopt_design = cordexch(AnzahlVariable,AnzahlVersuche,model,'tries',10);
X_dopt = [dopt_design(:,1)*30+50, dopt_design(:,2)*2100+2900];
M_dopt = Schleppmoment(X_dopt);
stats = regstats(M_dopt,X_dopt,model);
stats.tstat.pval
stats.adjrsquare
b_dopt = stats.tstat.beta

Mreg = b_dopt(1) + b_dopt(2)*Rval + b_dopt(3)*Nval + b_dopt(4)*Rval.*Nval +
b_dopt(5)*Nval.^2 + b_dopt(6)*Nval.^3+ b_dopt(7)*Nval.^4;
Mres = Mreg - Mval;
```

```
subplot(1,2,1);
hold on;
plot(Rval,Mres,'rx','linewidth',2);
hold off;

subplot(1,2,2);
hold on;
plot(Nval,Mres,'rx','linewidth',2);
hold off;
```

k) Zur statistischen Tolerierung wird zunächst symbolisch und anschließend numerisch die Empfindlichkeit E_x der einzelnen Einflussgrößen berechnet. Die Toleranzangaben werden in Standardabweichungen σ_x umgerechnet. Die Standardabweichung der Zielgröße ergibt sich aus

$$\sigma_m = \sqrt{E_p^2 \cdot \sigma_p^2 + E_T^2 \cdot \sigma_T^2 + E_\eta^2 \cdot \sigma_\eta^2 + E_{Vh}^2 \cdot \sigma_{Vh}^2 + E_n^2 \cdot \sigma_n^2}$$

Die sich daraus ergebende Standardabweichung wird mit dem Erweiterungsfaktor 6 multipliziert, da dies der Konfidenzzahl von $\gamma = 99.73$ % entspricht. Die Toleranz ergibt sich damit zu

$$T_{Stat} = 6 \cdot \sigma_m = 22.0381\,\text{kg/h}$$

```
% Statistische Tolerierung über den Grenzwertsatz
% Symbolische Berechnung der Ableitungen
syms p R T eta Vh n
m = p/R/T*eta*Vh*n;
E_p_s = diff(m,p);
E_T_s = diff(m,T);
E_eta_s = diff(m,eta);
E_Vh_s = diff(m,Vh);
E_n_s = diff(m,n);

% Übernehmen der Sollwerte und Toleranzen
p = 1013*100; sig_p = 0.01*p/6;
R = 287.1;
T = 323; sig_T = 2/6;
eta = 0.8; sig_eta = 0.04/sqrt(12);
Vh = 1900/1e6; sig_Vh = 10/sqrt(12)/1e6;
n = 2500*60; sig_n = 0.006*n/sqrt(12)

% Numerische Berechnung der Größen
m0 = p/R/T*eta*Vh*n
E_p = eval(E_p_s);
E_T = eval(E_T_s);
E_eta = eval(E_eta_s);
E_Vh = eval(E_Vh_s);
E_n = eval(E_n_s);
sig_m = sqrt(E_p^2*sig_p^2 + E_T^2*sig_T^2 + E_eta^2*sig_eta^2 + E_Vh^2*sig_Vh^2
+ E_n^2*sig_n^2);
T_stat = 6*sig_m
```

l) Zur statistischen Simulation werden Zufallszahlen erzeugt, die der angegebenen Toleranzverteilung entsprechen. Die aus den Zufallszahlen generierte Ausgangsgröße wird verwendet, um das Prognoseintervall für $\gamma = 99.73$ % zu bestimmen. Wegen der unbekannten Standardabweichung wird die t-Verteilung verwendet.

Es ergibt sich eine Toleranz von

$T_{Sim} = 22.0585$ kg/h

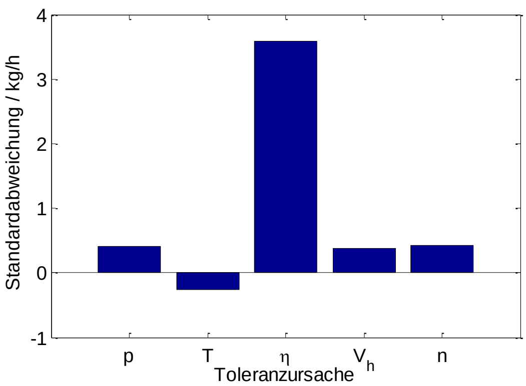
```
% Generieren von Zufallszahlen und Toleranzsimulation
N = 10000;
pv = normrnd(p,sig_p,N,1);
Tv = normrnd(T,sig_T,N,1);
etav = unifrnd(0.78,0.82,N,1);
Vhv = unifrnd(1895/1e6,1905/1e6,N,1);
nv = unifrnd(2500*0.997*60,2500*1.003*60,N,1);
mv = pv./R./Tv.*etav.*Vhv.*nv;

% Berechnung des Prognoseintervalls
c1 = tinv((1-0.9973)/2,N-1);
c2 = tinv((1+0.9973)/2, N-1);
T_sim = std(mv)*sqrt(1+1/N)*(c2 - c1)
```

m) Die beiden Toleranzen stimmen sehr gut überein. Das Ergebnis ist damit plausibel.

n) Zur Untersuchung, welche Toleranzursache dominiert, werden die Teiltoleranzen als Säulendiagramm dargestellt. Die Teiltoleranzen ergeben sich dabei aus dem Produkt von Empfindlichkeit und Standardabweichung.

$T_X = E_X \cdot \sigma_X$

Der variable Füllungsgrad ist die dominante Toleranzursache.




```
% Vergleich der Toleranzursachen
T_p = E_p*sig_p;
T_T = E_T*sig_T;
T_eta = E_eta*sig_eta;
T_Vh = E_Vh*sig_Vh;
T_n = E_n*sig_n;

% Grafische Darstelleung
names = {'p','T','\eta','V_h','n'};
Tol = [T_p T_T T_eta T_Vh T_n];
bar(1:length(Tol),Tol);
xlabel('Toleranzursache','FontWeight');
ylabel('Standardabweichung / kg/h');
```