

Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun

Heinrich-Heine-Universität Düsseldorf

22. Dezember 2010

1 Binomialtests

- Einseitiger unterer Binomialtest
- Zweiseitiger Binomialtest
- Beispiel BSE
- Normalapproximation
- Versuchsplanung
- Der p-Wert
- Data Snooping

$$\sum_{k=0}^c \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} > \alpha$$

$$\sum_{k=0}^c \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} > \alpha$$

Einseitiger unterer Binomialtest, Fortsetzung

- Der Annahmehereich ist

$$K_0 = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j \geq c \right\}$$

- Der kritische Bereich ist

$$K_1 = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j < c \right\}$$

- Das bedeutet: Die Nullhypothese wird abgelehnt, falls die Anzahl der Erfolge echt kleiner als c ist, andernfalls wird sie angenommen.

- Gegeben sind unabhängige $B(1, p)$ -verteilte Zufallsvariable $X_1, \dots, X_n$ mit unbekanntem p sowie ein Signifikanzniveau α
- Getestet wird die Nullhypothese $H_0 = \{p = p_0\}$ gegen die Alternative $H_1 = \{p \neq p_0\}$
- Der kritischen Wert c_1 und c_2 sind so zu wählen, dass

$$\sum_{k=0}^{c_1-1} \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{n-k} \leq \frac{\alpha}{2}$$

$$\sum_{k=0}^{c_1} \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} > \frac{\alpha}{2}$$

$$\sum_{k=0}^{c_2} \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{n-k} \geq 1 - \frac{\alpha}{2}$$

$$\sum_{k=0}^{c_2-1} \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{n-k} < 1 - \frac{\alpha}{2}$$

- Das bedeutet: Die Nullhypothese wird angenommen, wenn die Anzahl der Erfolge zwischen c_1 und c_2 liegt, andernfalls abgelehnt

- Bei 250 Würfeln eines Würfels fiel 55 mal eine Sechs. Kann man zu 98% sicher sein, dass der Würfel gezinkt ist?
- Sei p die unbekannte Wahrscheinlichkeit des Würfels für eine Sechs
- Zweiseitiger Binomialtest mit

Nullhypothese: $H_0 = \{p = \frac{1}{6}\}$

Alternative: $H_1 = \{p \neq \frac{1}{6}\}$

- Signifikanzniveau ist $\alpha = 0.02$

Tabelle von $\sum_{k=0}^n B_{250, 1/6}(k)$

r	p	$\frac{1}{6}$
20	0.	00005
21		00011
22		00024
23		00050
24		00100
25		00189
26		00343
27		00598
28		01005
29		01628
30		02546
31		03849
32		05632
33		07989
34		10997
35		14709
36		19143
37		24273
38		30023
39		36275

r	p	$\frac{1}{6}$
40	0.	42870
41		49627
42		56351
43		62856
44		68977
45		74581
46		79576
47		83912
48		87579
49		90603
50		93034
51		94941
52		96400
53		97491
54		98286
55		98853
56		99248
57		99517
58		99696
59		99812

Beispiel, Fortsetzung

- $c_1 = 28$ und $c_2 = 56$
- Die Nullhypothese kann zum Niveau $\alpha = 0.02$ abgelehnt werden, wenn höchstens 27 oder mindestens 57 Sechsen fallen
- Bei 55 Sechsen kann die Nullhypothese nicht abgelehnt werden

Effektive Fehlerwahrscheinlichkeit im Beispiel

- Der Fehler erster Art wird gemacht, wenn $p = p_0 = 1/6$ und ≤ 27 oder ≥ 57 Sechsen fallen

$$P(X \leq 27) = \sum_{k=0}^{27} \binom{250}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} = 0.00598$$

$$\begin{aligned} P(X \geq 57) &= 1 - \sum_{k=0}^{56} \binom{250}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} \\ &= 1 - 0.99248 = 0.00752 \end{aligned}$$

- Die effektive Fehlerwahrscheinlichkeit erster Art ist die Summe

$$\alpha_0 = 0.00598 + 0.00752 = 0.01350$$

Fehlerwahrscheinlichkeit zweiter Art im Beispiel

- Der Würfel ist gezinkt mit Wahrscheinlichkeit $p = \frac{1}{5}$, eine Sechs zu werfen
- Mit welcher Wahrscheinlichkeit wird das nicht entdeckt?
- X sei verteilt gemäß $B_{250, 1/5}$
- Gesucht ist

$$\begin{aligned}
 P(28 \leq X \leq 56) &= \sum_{k=0}^{56} \binom{250}{k} \cdot p^k \cdot (1-p)^{250-k} \\
 &\quad - \sum_{k=0}^{27} \binom{250}{k} \cdot p^k \cdot (1-p)^{250-k} \\
 &= 0.84786 - 0.00007 = 0.84779
 \end{aligned}$$

- Mit einer Wahrscheinlichkeit von fast 85% wird der gezinkte Würfel nicht entdeckt

Tabelle von $\sum_{k=0}^n B_{250, 1/5}(k)$

r	p	$\frac{1}{5}$
24	0.	00001
25		00002
26		00003
27		00007
28		00015
29		00030
30		00057
31		00106
32		00189
33		00327
34		00546
35		00884
36		01389
37		02120
38		03143
39		04534

r	p	$\frac{1}{5}$
40	0.	06368
41		08717
42		11639
43		15172
44		19328
45		24084
46		29383
47		35133
48		41213
49		47478
50		53775
51		59948
52		65855
53		71371
54		76402
55		80884
56		84786
57		88106
58		90868
59		93115

Beispiel BSE

- In einem Versuch wurden 275 Mäuse mit Milch von Kühen ernährt, welche an BSE erkrankt waren
- In keinem Fall wurde die Krankheit auf die Maus übertragen
- Was schließen wir daraus?

BSE, Fortsetzung

- Sei p die Wahrscheinlichkeit, dass BSE über die Nahrung von Kühen auf Mäuse übergeht
- Was ist die Nullhypothese?
- Die Ungefährlichkeit des Übertragungswegs soll gezeigt werden
- Die Nullhypothese ist also, dass der Übertragungsweg gefährlich ist
- Die Nullhypothese $H_0 = \{p > 0\}$ macht keinen Sinn
- Eine konkrete Schranke muss her
- Wähle $p_0 = 0.01$
- Es handelt sich also um einen einseitigen unteren Binomialtest. Als Signifikanzniveau wählen wir 5%

Einseitiger unterer Binomialtest zum Niveau α

- Der kritische Wert c ist so zu wählen, dass

$$\sum_{k=0}^{c-1} \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{n-k} \leq \alpha$$

$$\sum_{k=0}^c \binom{n}{k} \cdot p_0^k \cdot (1-p_0)^{n-k} > \alpha$$

- Der Annahmebereich ist

$$K_0 = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j \geq c \right\}$$

- Der kritische Bereich ist

$$K_1 = \left\{ (x_1, \dots, x_n) \mid \sum_{j=1}^n x_j < c \right\}$$

Tabelle von $\sum_{k=0}^r B_{275,p}(k)$

r	p	0.010	0.011	0.012	0.013	0.014
0	0.	06305	04775	03615	02737	02071
1		23819	19380	15691	12649	10157
2		48055	41635	35785	30535	25886
3		70332	64160	57994	51974	46210
4		85634	81196	76337	71175	65832
5		94011	91466	88411	84882	80934
6		97819	96606	95011	93006	90582
7		99297	98803	98091	97119	95847
8		99797	99622	99345	98933	98351
9		99947	99892	99796	99642	99406
10		99987	99972	99942	99890	99805
11		99997	99993	99985	99969	99941
12		99999	99999	99996	99992	99983
13				99999	99998	99996
14						99999

Beispiel BSE, Fortsetzung

- Zum Signifikanzniveau $\alpha = 0.05$ ist der kritische Bereich leer
- Mit 275 Experimenten ist daher keine Aussage möglich
- Wir können zu diesem Signifikanzniveau aber die folgende, veränderte Nullhypothese ablehnen

$$H_0 = \{p \geq 0.011\}$$

- Mit 95%-tiger Sicherheit können wir sagen, dass das Risiko einer Übertragung von BSE auf Mäuse kleiner als 1.1% ist

Normalapproximation

- Bei großen Stichprobenumfängen verwende Normalapproximation zur Bestimmung des kritischen Werts
- Beispiel: einseitiger oberer Binomialtest zum Niveau α
- Dann ist c die kleinste ganze Zahl, für die

$$\sum_{k=0}^c \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} \geq 1 - \alpha$$

- Ersetze die kumulierte Binomialverteilung durch die Normalapproximation

Normalapproximation, Erinnerung

- Die Zufallsvariable X sei $B(n, p)$ -verteilt mit

$$n \cdot p \cdot (1 - p) > 9$$

- Dann gilt näherungsweise für natürliche Zahlen $a < b$

$$P(a \leq X \leq b) \cong \Phi\left(\frac{b + \frac{1}{2} - n \cdot p}{\sqrt{n \cdot p \cdot (1 - p)}}\right) - \Phi\left(\frac{a - \frac{1}{2} - n \cdot p}{\sqrt{n \cdot p \cdot (1 - p)}}\right)$$

- Wenn $a = 0$ oder $b = n$ ist, braucht man nur einen Term

$$P(a \leq X) \cong 1 - \Phi\left(\frac{a - \frac{1}{2} - n \cdot p}{\sqrt{n \cdot p \cdot (1 - p)}}\right)$$

$$P(X \leq b) \cong \Phi\left(\frac{b + \frac{1}{2} - n \cdot p}{\sqrt{n \cdot p \cdot (1 - p)}}\right)$$

Normalapproximation

- Gesucht ist also die Lösung c der Gleichung

$$P(X \leq c) \cong \Phi\left(\frac{c + 1/2 - n \cdot p_0}{\sqrt{n \cdot p_0 \cdot (1 - p_0)}}\right) = 1 - \alpha$$

- Die so gefundene Zahl ist in der Regel nicht ganz. Man wählt dann als c die nächstgelegene ganze Zahl
- Die Voraussetzung für die Anwendbarkeit der Normalapproximation ist zu beachten

$$n \cdot p_0 \cdot (1 - p_0) > 9$$

Beispiel: Saatgut

- Die Keimfähigkeit eines Saatguts betrage p mit unbekanntem $p \in [0, 1]$. Bei $p \leq 0.75$ lohnt es nicht, das Saatgut auszubringen
- Um die Qualität des Saatguts zu testen, werden $n = 200$ Körner ausgesät. Nach gegebener Zeit wird die Zahl der gekeimten Körner bestimmt
- Wie lässt sich zum Signifikanzniveau $\alpha = 0.05$ sicher stellen, dass $p > 0.75$ gilt?
- Die Nullhypothese ist bei dieser Fragestellung, dass das Saatgut nichts taugt, dass also $p \leq p_0$, wobei $p_0 = 0.75$
- Es handelt sich um einen einseitigen oberen Binomialtest

Saatgut, Fortsetzung

- Prüfe die Anwendbarkeit der Normalapproximation:

$$n \cdot p_0 \cdot (1 - p_0) = 200 \cdot 0.75 \cdot 0.25 = 37.5$$

- $n \cdot p = 150$ und $\sqrt{37.5} = 6.12$, also bestimme c durch

$$\Phi\left(\frac{c + 1/2 - 150}{6.12}\right) = 1 - 0.05$$

- D. h.

$$\Phi\left(\frac{c - 149.5}{6.12}\right) = 0.95$$

- Das 0.95-Quantil der Standard-Normalverteilung ist 1.645

$$\frac{c - 149.5}{6.12} = 1.645$$

Saatgut, Fortsetzung

- Löse Gleichung nach c auf

$$c = 149.5 + 6.12 \cdot 1.645 = 159.6$$

- Die nächstgelegene ganze Zahl ist $c = 160$
- Bei 161 oder mehr Keimerfolgen wird die Nullhypothese abgelehnt und das Saatgut akzeptiert

Versuchsplanung

- Die Normalapproximation kann zur Versuchsplanung benutzt werden
- Ein Beispiel hatten wir in der Vorlesung vom 24. November betrachtet

Normalapproximation zur Bestimmung eines Stichprobenumfangs

- Zwei Würfel: Einer ist fair, einer gezinkt; bei dem gezinkten ist die Wahrscheinlichkeit einer 6 gleich $\frac{1}{5}$
- Will herausbekommen, welchen ich in der Hand habe
- Die Anzahl n der nötigen Würfe soll bestimmt werden
- Wahrscheinlichkeit, den fairen Würfel für gezinkt zu halten, soll höchstens 1% betragen
- Wahrscheinlichkeit, den gezinkten Würfel für fair zu halten, soll höchstens 5% betragen

Interpretation des Beispiels als Binomialtest

- Im Prinzip hatte ich damals einen einseitigen oberen Binomialtest vorgestellt

$$H_0 : p \leq \frac{1}{6}$$

$$H_1 : p > \frac{1}{6}$$

- Als Signifikanzniveau hatte ich 0.01 vorgeschrieben
- Bei $p = \frac{1}{5}$ sollte die Fehlerwahrscheinlichkeit zweiter Art bei 0.05 liegen
- Normalapproximation ergab $n = 2093$
- Bei 389 oder mehr Sechsen ist zum Signifikanzniveau der Nachweis geführt, dass der Würfel gezinkt ist
- Der kritische Wert war also 388

Der p -Wert, Idee

- Gelegentlich ist nicht ganz klar, welches Signifikanzniveau α angemessen ist
- Man macht dann seinen Versuch und gibt das beste Signifikanzniveau an, für das die Nullhypothese abgelehnt werden kann
- Der p -Wert eines Experiments ist das kleinste Signifikanzniveau, zu dem die Beobachtungsdaten die Annahme der Alternative rechtfertigen

Der p -Wert

- Es sei Θ eine Menge von Parametern. Zu jedem $\theta \in \Theta$ gebe es eine Verteilung P_θ . Die Zufallsvariablen $X_1, \dots, X_n$ seien unabhängig und alle nach demselben P_θ verteilt
- $H_0 \subset \Theta$ bezeichne die Nullhypothese
- Für jedes mögliche Signifikanzniveau α sei eine Testvorschrift festgelegt. Ihr Annahmebereich sei $K_{0,\alpha}$ und ihr kritischer Bereich sei $K_{1,\alpha}$
- Es seien $x = (x_1, \dots, x_n)$ die Beobachtungsdaten
- Das kleinste α mit $x \in K_{1,\alpha}$ bezeichnet man als p -Wert der Beobachtungsdaten
- p ist also das kleinste Signifikanzniveau, zu dem die Nullhypothese noch abgelehnt werden kann
- Das bedeutet, dass die Beobachtungsdaten x genau dann zur Ablehnung der Nullhypothese führen, wenn ihr p -Wert höchstens α beträgt

Der p -Wert, Beispiel

- Im Beispiel "Saatgut" haben 158 Körner gekeimt. Was ist der p -Wert?
- Vorüberlegung: Jedenfalls $p > 0.05$, denn für $\alpha = 0.05$ haben wir bereits ausgerechnet, dass der kritische Bereich erst bei 161 Keimungen beginnt.
- Beim oberen Binomialtest ist c die kleinste Zahl, für die

$$\sum_{k=0}^c \binom{n}{k} \cdot p_0^k \cdot (1 - p_0)^{n-k} \geq 1 - \alpha$$

- Bei der Bestimmung des p -Werts kennen wir das c . Das zugehörige α ist der p -Wert
- Im Beispiel muss $c = 157$ sein oder noch kleiner, damit 158 im kritischen Bereich ist

Beispiel, Fortsetzung

- Das α , das zu $c = 157$ gehört, ist der p -Wert
- Also

$$1 - p = \sum_{k=0}^{157} \binom{200}{k} \cdot 0.75^k \cdot 0.25^{200-k}$$

- Verwende Normalapproximation

$$\begin{aligned} \sum_{k=0}^{157} \binom{200}{k} \cdot 0.75^k \cdot 0.25^{200-k} \\ \cong \Phi \left(\frac{157 + 1/2 - 200 \cdot 0.75}{\sqrt{200 \cdot 0.75 \cdot (1 - 0.75)}} \right) \\ = \Phi \left(\frac{7.5}{6.12} \right) = \Phi(1.23) = 0.890651 \end{aligned}$$

- Der p -Wert beträgt 11%

Data Snooping

- “Snooping” = “Schnüffeln”
- Data Snooping bedeutet, dass man den Test für dieselben Daten rechnet, die man auch für die Formulierung der Hypothese benutzt hat
- Die nächste Folie stammt aus einem schlechten Buch (und wird daher am Netz nicht gezeigt)