PARTICLE AWARECON „STATISTIC- A LIFE SLOT“ THEORIE
Anzahl der Geräte schätzen

- Jeder Knoten sendet während des Alife slots synchron Bits mit bestimmten Wahrscheinlichkeiten
- Jeder Knoten empfängt zur gleichen Zeit auch die gesendeten Sachen
- Die Sendewahrscheinlichkeiten sind wie folgt verteilt
-

Bitnummer	1	2	3	4	5	6	7	8
Sendewahrscheinlichkeit	1	1/2	1/4	1/8	1/16	1/32	1/64	1/128

- Ziel ist es, im Empfänger die Anzahl der Geräte über einen gewissen Zeitraum zu schätzen
- Dafür werden Empfangene Daten aus dem AliveSlot gespeichert und ausgewertet

Festlegung von Bezeichnern:

x: Bitposition
k: Anzahl der anwesenden Geräte
$\tilde{k}$: geschätzte Zahl
$\tilde{k}_x$: geschätzte Zahl der anwesenden Geräte aus Schätzung von Bitpos.x
r_{rx}: relative Häufigkeit von empfangenen Einsen auf Bitnummer x
n: Anzahl der Beobachtungsslots
p_{ax}: Apriori Sendewahrscheinlichkeit (Senden einer eins) auf Bitnummer x
p_{rx}: Empfangswahrscheinlichkeit (Empfangen einer eins) auf Bitnummer x
$\tilde{p}_{rx}$: geschätzte Empfangswahrscheinlichkeit auf Bitnummer x
a_x: empfangene Einsen auf Bitnummer x

Szenario: k Geräte anwesend

P(Empfänger empfängt Bit auf Position x; k Geräte anwesend)

$$p_{rx} = 1 - (1 - p_{ax})^k$$

Zufallsvariable $A_{x,n}$ gibt die Anzahl von empfangenen Einsen an, bei einer Beobachtungszeit von n Slots; (hängt ab von Parametern x, n, p_{rx}) Realisierung)

$$P(A_{x,n} = a_x) = p_{rx}^{a_x} \cdot (1 - p_{rx})^{n-a_x} \cdot \binom{n}{a_x}$$; Binomialverteilung!

- $E(A_{x,n}) = n \cdot p_{rx}$
- $Var(A_{x,n}) = n \cdot p_{rx} \cdot (1 - p_{rx})$

AUFGABENSTELLUNG 1

Wie kann nun der Empfänger auf die unbekannte Anzahl von Geräten schliessen?

LÖSUNG 1

Schätzaufgabe! => Maximum Likelihood Schätzer
Likelihood Funktion (gibt Wahrscheinlichkeit an, dass der Fall $A_{x,n} = a_x \,|_{p_{rx}}$ eintritt):

$$L_k(p_{rx}) = P(A_{x,n} = a_x) = p_{rx}{}^{a_x} \cdot (1-p_{rx})^{n-a_x} \cdot \binom{n}{a_x} \text{ (k fest, } p_{rx} \text{ Parameter)}$$

Maximum suchen:

$$\frac{\partial}{\partial p_{rx}} L_k(p_{rx}) \overset{!}{=} 0 = p_{rx}{}^{a_x-1} \cdot (1-p_{rx})^{n-a_x-1} \cdot (a_x(1-p_{rx}) - (n-a_x)p_{rx}) \cdot \binom{n}{a_x} = 0$$

Lösung $\tilde{p}_{rx} = \frac{a_x}{n}$; also die relative Häufigkeit (Überraschung ;))

Zufallsvariable $R_{x,n} = \frac{A_{x,n}}{n}$ für die Schätzung $\tilde{p}_{rx}$ gibt die relative Häufigkeit

- $E(R_{x,n}) = \frac{1}{n} \cdot E(A_{x,n}) = \frac{n \cdot p_{rx}}{n} = p_{rx}$
- $Var(R_{x,n}) = \frac{1}{n^2} \cdot Var(A_{x,n}) = \frac{n \cdot p_{rx} \cdot (1-p_{rx})}{n^2} = \frac{p_{rx} \cdot (1-p_{rx})}{n}$
- für $n \to \infty$: Varianz strebt gegen null. Erwartungswert ist p_{rx};
- Schätzung ist erwartungstreu!!

Der Empfänger schätzt die Empfangswahrscheinlichkeit $\tilde{p}_{rx}$ und kann dann die Zahl k der Geräte mit

$$\tilde{k}_x = \frac{\ln(1-\tilde{p}_{rx})}{\ln(1-p_{ax})}$$

ausrechnen.
Dies macht der Empfänger für alle Bitpositionen x ein $\tilde{k}_x \quad x = 2..8$ k. Für $\tilde{p}_{rx}$, die null oder eins sind, ist keine sinnvolle Berechnung möglich.

PROBLEM:
Der Schätzwert für $\tilde{p}_{rx}$ ist schlicht die relative Häufigkeit der „Eins“ auf der Bitposition x. ($\tilde{p}_{rx}$ ist erwartungstreu). Hier können also – besonders für kleine Beobachtungsintervalle – massiv Fehler auftreten. Somit ist auch die Schätzung für die Anzahl der Geräte $\tilde{k}_x$ fehlerhaft. Man muss sich also über den Fehler, den man macht, im Klaren werden!

AUFGABENSTELLUNG 2

Der gefundene Schätzwert $\tilde{p}_{rx}$ soll qualitativ beurteilt werden.

LÖSUNG 2

Angabe einer Sicherheit (Wahrscheinlichkeit), dass die geschätzte Wahrscheinlichkeit $\tilde{p}_{rx}$ eine gewisse Nähe zur tatsächlichen p_{rx} hat.
Nun sind verschiedene Fragestellungen möglich:

1. Vorgabe eines Vertrauensintervall $k \in [\tilde{k} - (\tilde{k} \cdot dk); \tilde{k} + (\tilde{k} \cdot dk)]$ (gleichwertig: $p_{rx} \in [\tilde{p}_{rx} - (\tilde{p}_{rx} \cdot d\tilde{p}_{rx}); \tilde{p}_{rx} + (\tilde{p}_{rx} \cdot d\tilde{p}_{rx})]$) und einer Sicherheit a (typ. >95%) => Frage nach den nötigen Stichproben n
2. Vorgabe einer Sicherheit a und der Stichprobenanzahl n => Frage nach dem Vertrauensintervall, das den Schätzwert $\tilde{k}_x$ umgibt
3. Vorgabe des Vertrauensintervalls, der Stichprobenanzahl n => Frage nach der Sicherheit a, mit dem das Vertrauensintervall gültig ist.

Weitere Betrachtung der Fragestellung 2:

Die erste Abschätzung gelingt mit der Ungleichung von Tschebyscheff

$$P(|X - E(X)| \geq e) \leq \frac{1}{e^2} \cdot Var(X)$$

eingesetzt für die zufällige, relative Trefferhäufigkeit:

$$P(|R_{x,n} - p_{rx}| \geq e) \leq \frac{1}{e^2} \cdot \frac{p_{rx} \cdot (1 - p_{rx})}{n} \leq \frac{1}{4e^2 n} \leq 1 - a$$

Beispiel:

$n = 30; a = 95\%; x = 5; a_5 = 6$ (30 Slots, Bitnummer 5 beobachtet, 6mal die Eins gesehen)

Schätzwerte: $\tilde{p}_{r5} = \frac{a_5}{n} = 0.2;\ \tilde{k}_5 = 3;$

Damit ergibt sich die rechte Seite zu

$$\frac{1}{4e^2 n} = 1 - 95\% \Leftrightarrow e = 0.41$$

mit $\tilde{p}_{r5}$=0.2 gilt jetzt die Aussage:	$P(p_{r5} \in [0; 0.61]) \geq 0.95$
Für die Gerätezahl gilt dann analog mit Gl bla	$P(k \in [0; 15]) \geq 0.95$
Bei gleichen Bedingungen, aber $n = 300$ gilt:	$P(k \in [1; 7]) \geq 0.95$

Die Aussage über die Geräteanzahl ist also bei geringen Stichproben äußerst ungenau. Dies soll nun verbessert werden:

Die prinzipielle Idee ist, mit den Vertrauensintervallgrenzen p_{unten}, p_{oben} ungültige p auszuschließen. Dafür folgende Überlegung. Gegeben sei der Stichprobenumfang n und die Trefferzahl a_x. Es lässt sich nun ein p_{rx} bestimmen, für das gilt, dass die Wahrscheinlichkeit für höchstens a_x hinreichend klein ($\leq b$) wird:

$$P(a \leq a_x) = \sum_{i=0}^{a_x} \binom{n}{i} p_{rx}{}^{i} (1 - p_{rx})^{n-1} \leq b; \quad b = \frac{1-a}{2}.$$

Durch die Monotonie gilt die Ungleichung erst recht für größere p_{rx}. Somit ist eine obere Grenze für p_{rx}(!!!) gefunden. Analog verfährt man für die untere Grenze:

$$\tilde{p}_{rx} < p_{unten} : P(a \geq a_x) = \sum_{i=a_x}^{n} \binom{n}{i} p_{rx}^{\ i} (1 - p_{rx})^{n-1} \leq b$$

Wenn beide Aussagen mit $1 - \frac{1-a}{2}$ gültig sind, dann ist die Aussage

$P(p_{rx} \in [p_{unten}; p_{oben}]) > a$ mit eben der Sicherheit a gültig. Bei Vorgabe der gewünschten Sicherheit a und Angabe von n und den Treffern a_x kann man das Vertrauensintervall $[p_{unten}; p_{oben}]$ angeben.

Die Vertrauensintervallgrenzen können dann mit Hilfe von $p_{rx} = 1 - (1 - p_{ax})^k$ auf k abgebildet werden, was wiederum auf ein Vertrauensintervall für gefundene Geräte führt.