

Übungsserie 4

1. In einer Schokoladefabrik werden pro Stunde im Schnitt 6 Kirschstängeli produziert, die als Ausschuss weggeworfen werden müssen. Sinnvollerweise modellieren wir den stündlichen Ausschuss mit einer Poissonverteilung.

- a) Berechnen Sie die Wahrscheinlichkeit $P(X=k)$ für k von 0 bis 20. Wie gross ist die Wahrscheinlichkeit, dass 10 Kirschstängeli weggeworfen werden müssen? Wie gross ist die Wahrscheinlichkeit dass 10 *oder mehr* weggeworfen werden müssen?

Erzeugen Sie eine leere Datentabelle mit **File / New / Data**. In die erste Zelle der ersten Spalte klicken, die Spalte in **Data / Variable Properties ...** mit **K** benennen und die Anzahl Dezimalstellen auf **0** setzen. Dann mit **Data / Fill worksheet** die Anzahl benötigter Zeilen auf 21 setzen. Nachher im Menü **Data / Transform / Let** können Sie nun **K** die Werte 0 bis 20 zuweisen (**K = CASE-1**) und dann im selben Menü die zugehörige Wahrscheinlichkeit als neue Variable berechnen: Die Funktion **PDF(k,λ)** gibt als Resultat $P(X=k)$ unter der Annahme, dass $X \sim \text{Poisson}(\lambda)$. Setzen Sie für k und λ die richtigen Werte ein.

- b) Sie wollen kurz mal überprüfen, ob die Kirschstängelmaschine regulär arbeitet und schauen zu diesem Zweck den Ausschuss der letzten Stunde an. Ab welchem Ausschuss werden Sie vermuten, dass da etwas nicht mehr stimmt? (Wählen Sie $\alpha = 5\%$ als Niveau)

Tragen Sie mittels eines Stabdiagramms die Wahrscheinlichkeit gegen die Anzahl auf und beantworten Sie die Frage mit Hilfe dieser Grafik und der Tabelle aus **a**).

2. Zwei unterschiedlich hergestellte Käsesorten wurden auf ihre Bitterkeit hin untersucht. In der Datei **kaese.syd** finden Sie eine Tabelle der Bitterkeitswerte, die den beiden Käsesorten von 20 PrüferInnen zugeteilt wurden. Wir wollen nun schauen, wie sehr sich die Sorten unterscheiden.

- a) Berechnen Sie für jede Person die Differenz der beiden Bewertungen (**DIFF=A-B**). Erstellen Sie ein Stabdiagramm von **DIFF** und beschreiben Sie es.
- b) Berechnen Sie den Mittelwert von **DIFF**. Was sagt er aus?
- c) Geben Sie die 90%-, 95%- und 99%-Vertrauensgrenzen für den Mittelwert μ der Gesamtheit der Differenzen an. Weshalb wird der Bereich innerhalb der Grenzen grösser, wenn das Vertrauensniveau erhöht wird?

Im Menü **Statistics / Descriptive Statistics / Basic Statistics** können Sie sich das Konfidenzintervall angeben lassen, indem Sie **CI** (Abkürzung für "confidence interval") auswählen und die gewünschte Prozentzahl angeben.

Achtung: Dieses Konfidenzintervall wird mit der Annahme berechnet, die Daten seien normalverteilt. Wenn diese Annahme nicht stimmt, so sind die Resultate mit Vorsicht zu interpretieren!

- d) Entscheiden Sie aufgrund der obigen Berechnungen, ob sich die beiden Käsesorten in der Bitterkeit systematisch unterscheiden, oder ob die Unterschiede in den Bewertungen rein zufällig sein können. Begründen Sie. (Liegt 0 im Vertrauensintervall? Was bedeutet das?)

3. Beim Abfüllen von Fruchtsaft in Literflaschen kommt es der Sana Saft AG sehr auf das genaue Einhalten der Füllmenge an, denn die Kunden reklamieren schnell, wenn die Füllmenge einen Liter unterschreitet, und andererseits möchte die Sana Saft AG natürlich Ertragseinbussen durch zu grosszügig gefüllte Flaschen vermeiden. Der Fruchtsaftinhalt einer Literflasche ist normalverteilt mit Erwartungswert $\mu = 1$ L und Streuung $\sigma = 0.004$ L.
- Weicht der gemessene Inhalt einer Flasche um mehr als 0.01 L vom Sollwert μ ab, so wird die Flasche nicht akzeptiert. Wie gross ist der Ausschussanteil in Prozent?
 - Bestimmen Sie eine Zahl d so, dass der gemessene Inhalt von 95% aller Flaschen zwischen $(1-d)$ L und $(1+d)$ L liegt (runden Sie auf 3 Stellen nach dem Komma).
4. Die komplexen Kohlenhydrate stellen einen wichtigen Teil für die Kalorienzufuhr in einer balancierten Diät dar. Bei einer Studie wurde ein Sortiment von 81 Getreideflocken untersucht. Im Mittel erhielt man pro 100 g Flocken 14.3 g komplexe Kohlenhydrate. Die Standardabweichung der einzelnen Bestimmungen betrug 3.9 g.
- Berechnen Sie den Standardfehler des Mittelwertes des Kohlenhydratengehaltes (pro 100 g Flocken).
 - Angenommen die Daten seien normalverteilt, bestimmen Sie ein 95%-Konfidenzintervall für den wahren Gehalt an komplexen Kohlenhydraten (pro 100 g Flocken).
 - Auf allen Verpackungen ist ein Inhalt von 15 g angegeben. Ist dieser Wert mit den Daten verträglich (auf dem 5% Niveau)? Begründen Sie Ihre Antwort.

Punkteverteilung:

- Aufgabe 1:** a) (1 Punkt) $P(X = 10)$ und $P(X \geq 10)$.
b) (1 Punkt) Kritischer Wert mit Begründung.
- Aufgabe 2:** a) (1 Punkt) Kurze Beschreibung des Stabdiagramms.
b) (1 Punkt) Mittelwert μ mit ganz kurzer Interpretation.
c) (1 Punkt) Alle drei Konfidenzintervalle.
d) (1 Punkt) Entscheid und Begründung.
- Aufgabe 3:** a) (1 Punkt) Wahrscheinlichkeit.
b) (1 Punkt) Berechnung von d .
- Aufgabe 4:** a) und b) (1 Punkt) Standardfehler und Konfidenzintervall.
c) (1 Punkt) Entscheid und Begründung.

Abgabe: 1 Woche nach der Übungsstunde in der Vorlesung.

Präsenz: Montag, 12:00-13:00, LEO C12.1.