Package 'msamp'

July 23, 2025

Title Estimate Sample Size to Detect Bacterial Contamination in a
Product Lot

Version 1.0.0

Description Estimates the sample size needed to detect microbial contamination
in a lot with a user-specified detection probability and user-specified analytical sensitivity.
Various patterns of microbial contamination are accounted for: homogeneous (Poisson),
heterogeneous (Poisson-Gamma) or localized(Zero-inflated Poisson).
Ida Jongenburger et al. (2010) <doi:10.1016/j.foodcont.2012.02.004>
``Impact of microbial distributions on food safety".
Leroy Simon (1963) <doi:10.1017/S0515036100001975>
``Casualty Actuarial Society - The Negative Binomial and Poisson Distributions Compared".

License Unlimited

Encoding UTF-8

RoxygenNote 7.2.1

VignetteBuilder knitr

Suggests knitr, rmarkdown

Imports graphics, stats, grDevices

NeedsCompilation no

Author Martine Ferguson [aut, cre] (ORCID:
<https://orcid.org/0000-0003-4479-3674>)

Maintainer Martine Ferguson <martine.ferguson@fda.hhs.gov>

Repository CRAN

Date/Publication 2022-10-17 12:00:05 UTC

Contents

---

n   *Calculate the sample size necessary to detect contamination above tar-*
                                         get level

Description

The n() function calculates the sample size,n, necessary to detect contamination above a target
level, G, in a product lot, where the probability of a single sample unit being above the target level
is calculated by the msamp function p(). The total cost, cost_tot, associated with sample size is also
output.

Usage

```
n(
   C,
   w,
   G,
   Sens,
   D = c("homogeneous", "heterogeneous", "localized"),
   r = NULL,
   f = NULL,
   prob_det = 0.9,
   samp_dollar,
   lot_dollar
)
```

Arguments

C	suspected lot contamination (CFU/g)
w	weight of single sample unit (g)
G	target value to detect (CFU/g)
Sens	sensitivity of the analytical test (%)
D	distribution of the bacteria in the product lot: "homogeneous", "heterogeneous", or "localized"
r	for the heterogeneous case only, the degree of heterogeneity. $r > 0$
f	for the localized case, r is further specified. $0 < f < 1$
prob_det	desired probability of detecting bacterial contamination above the target level in the product lot. Set to 0.9 by default
samp_dollar	cost per sample unit in $
lot_dollar	fixed cost (if any) of sampling the lot in $

Details

Refer to vignette for details.

Value

A list containing:

- **n**: The sample size
- **p**: The probability of a single sample unit being contaminated above the target level
- **cost_tot**: The total cost associated with sampling of *n* samples

Examples

```
#A sample of 25 grams (w=25) is collected and analyzed using an analytical
#test with sensitivity of 90% (Sens=.9), to detect at least 5 CFU's/g (G=5).
#The suspected or postulated level of contamination in the lot is 4 CFU's/g (C=4).
#The desired probability of picking at least one sample unit contaminated above the target
#level is 0.9 (prob_det=0.9), the cost of a single sampling unit is $100 (samp_dollar=100),
#and the fixed cost for sampling the entire lot is $200 (lot_dollar=200).

#homogeneous case
n(C=4,w=25,G=5,Sens=.9,D="homogeneous",r=NULL,f=NULL,prob_det=0.9,samp_dollar=100,lot_dollar=200)
# n=376, total cost=$37,722
#heterogeneous case
n(C=4,w=25,G=5,Sens=.9,D="heterogeneous",r=10,f=NULL,prob_det=0.9,samp_dollar=100,lot_dollar=200)
# n=12, total cost=$1,319
#localized case
n(C=4,w=25,G=5,Sens=.9,D="localized",r=NULL,f=.3,prob_det=0.9,samp_dollar=100,lot_dollar=200)
# n=1,254 , total cost=$125,541
```

---

p *Calculate the probability of a single sample being contaminated*

Description

The p() function calculates the probability of a single sample unit, with weight w, and postulated contamination, C, being contaminated above a target level, G.

Usage

```
p(
  C,
  w,
  G,
  Sens,
  D = c("homogeneous", "heterogeneous", "localized"),
  r = NULL,
  f = NULL
)
```

Arguments

C	suspected lot contamination (CFU/g)
w	weight of single sample unit (g)
G	target value to detect (CFU/g)
Sens	sensitivity of the analytical test (%)
D	distribution of the bacteria in the product lot: "homogeneous", "heterogeneous", or "localized"
r	for the heterogeneous case only, the degree of heterogeneity. $r > 0$
f	for the localized case, r is further specified. $0 < f < 1$

Details

Refer to vignette for details.

Value

A numeric value: the probability of a single sample unit being contaminated above target level.

Examples

```
#A sample of 25 grams (w=25) is collected and analyzed using an analytical
#test with sensitivity of 90% (Sens=.9), to detect at least 5 CFU's/g (G=5).
#The suspected or postulated level of contamination in the lot is 4 CFU's/g (C=4)

#homogeneous case
p(C=4,w=25,G=5,Sens=.9,D="homogeneous",r=NULL,f=NULL)
# 0.006117884
#heterogeneous case-- dispersion, r, is postulated as 2
p(C=4,w=25,G=5,Sens=.9,D="heterogeneous",r=2,f=NULL)
# 0.2576463
#localized case -- 30% of the lot is postulated to be contaminated
p(C=4,w=25,G=5,Sens=.9,D="localized",r=NULL,f=.3)
# 0.001835365
```

---

plotn	*Plots the relation between the probability of detection and the sample size, n*

Description

The plotn() function examines the effect of increasing the probability of detection on the sample size, n, where the probability of a single sample unit being contaminated above the target limit is calculated from the msamp function p()

Usage

```
plotn(
  C,
  w,
  G,
  Sens,
  D = c("homogeneous", "heterogeneous", "localized"),
  r = NULL,
  f = NULL
)
```

Arguments

C	suspected lot contamination (CFU/g)
w	weight of single sample unit (g)
G	target value to detect (CFU/g)
Sens	sensitivity of the analytical test (%)
D	distribution of the bacteria in the product lot: "homogeneous", "heterogeneous", or "localized"
r	for the heterogeneous case only, the degree of heterogeneity. $r > 0$
f	for the localized case, r is further specified. $0 < f < 1$

Details

Refer to vignette for details.

Value

A plot, of recordedplot class. The probability of detection is on the y-axis and the sample size n is on the x-axis. Overlaid at intersecting red dashed lines is the sample size for probability of detection (prob_det) = 0.9.

Index