Discrete probability distributions II

Deependra Dhakal

https://rookie.rbind.io

Academic year 2019-2020

Poisson distribution

Definition

A random variable X is said to follow Poisson distribution if it assumes only non-negative values and its probability mass function is given by:

$$P(k,\lambda) = P(X=k) = \begin{cases} e^{-\lambda}\frac{\lambda^k}{k!}; x = 0,1,2,...,n; \lambda > 0 \\ 0, \ \text{otherwise} \end{cases}$$

Here, λ is known as the parameter of the distribution. We shall use the notation $X \sim p(\lambda)$, to denote that X is a poisson variate with parameter λ.

A poisson distribution is a limiting version of the binomial distribution, where n becomes large and np approaches some λ, which is the mean value.

The poisson distribution can be used for the number of events in other specified intervals such as distance, area or volume. Examples that may follow a Poisson include the number of phone calls received by a call center per hour and the number of decay events per second from a radioactive source.

Problem 1

The average number of goals in a World Cup football match is 2.5.

Probability of 4 goals in a match can be calculated as:

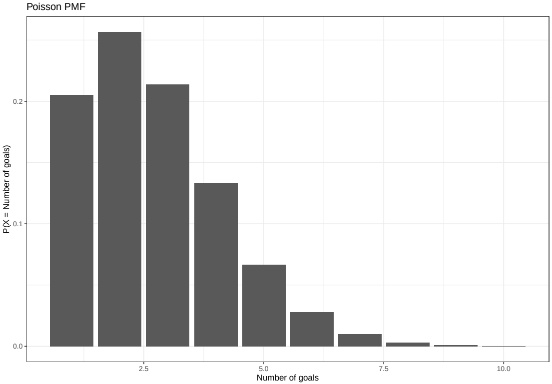
```
## [1] 0.133601885781085
```

This can be accomplished using built-in function

```
## [1] 0.133601885781085
```

Problem

Find probabilities of occurance of 1:10 goals and plot the poisson probability distribution



Negative binomial distribution

Definition

A random variable X is said to follow a negative binomial distribution with parameters r and p if its probability mass function is given by:

$$P(X=x)=p(x)=\begin{cases}\binom{x+r-1}{r-1}p^r q^x & x=0,1,2\ldots\\ 0, & \text{otherwise}\end{cases}$$

Geometric distribution

Definition

A random variable X is said to have a geometric distribution if it assumes only non-negative values and its probability mass function is given by:

$$P(X = x) = \begin{cases} q^x p; & x = 0, 1, 2 \ldots; 0 < p \leq 1; q = 1 - p \\ 0, & \text{otherwise} \end{cases}$$

Hypergeometric distribution

Definition

A random variable X is said follow the hypergeometric distribution with its parameters N, M and n if it assumes only non-negative values and its pmf is given by:

$$P(X = k) = h(k; N, M, n) \begin{cases} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}; k = 0, 1, 2, ..., min(n, M). \\ 0, \qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

Multinomial distribution

Meaning

This distriubtion can be regarded as a generalization of Binomial distribution.

When there are more than two mutually exclusive outcomes of a trial, the observations lead to multinomial distribution. Suppose $E_1, E_2, ..., E_k$ are k mutually exclusive and exhaustive outcomes of a trial with respective probabilities $p_1, p_2, ...p_k$.

The probability that E_1 occurs x_1 times, E_2, occurs x_2 times . . . and E_k, occurs x_k times in n independent observations, is given by $p(x_1, x_2, ..., x_k) = cp_1^{x_1} p_2^{x_2} ... p_k^{x_k}$, where $\sum x_i = n$ and c is the number of permutation of the events $E_1, E_2, ..., E_k$.

To determine c, we have to find the number of permutations of n objects of which x_1 are of one kind, x_2 of another kind, . . . , x_k of the kth kind, which is given by:

$$c = \frac{n!}{x_1!x_2!...x_k!}$$

$$\text{Hence} \quad p(x_1, x_2, ..., x_k) = \frac{n!}{x_1!x_2!...x_k!} p_1^{x_1} p_2^{x_2} ... p_k^{x_k}, 0 \leq x_i \leq n$$

$$= \frac{n!}{\prod_{i=1}^{k} x!} \prod_{i=1}^{k} p_i^{x_i}; \sum_{i=1}^{k} x_i = n$$

Which is the required probability function of the multinomial distribution. It is so called since the above expression is the general term in the multinomial expansion:

$$(p_1 + p_2 + ... + p_k)^n, \sum_{i=1}^{k} p_i = 1$$

Power series distribution

Definition

A discrete r.v. X is said to follow a generalized power series distribution (g.p.s.d), if its probability mass function is given by:

$$P(X = x) = \begin{cases} \frac{a_x \theta^x}{f(\theta)}; & x = 0, 1, 2 \ldots; a_x \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

Where $f(\theta)$is a generating function i.e.,

$$f(\theta) = \sum_{x \in s} a_x \theta^x, \theta \geq 0$$

So that $f(\theta)$ is positive, finite and differentiable and S is a non-empty countable subset of non-negative integers.

Normal distribution

Normal density

The `dnorm(x, mean = 0, sd = 1, log = FALSE)` function simply calculates the result for the value plugged into the probability density distribution or probability mass function if it is a discrete distribution.

So for the normal distribution with $mean = 0$, $sd = 1$, we have

$$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

If we plug $x = 2$ inside the pdf, we have

```
## [1] 0.0539909665131881
```

Normal distribution function

`pnorm(q, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)` returns the probabality of $p(X \leq x)$ by default. If we set `low.tail = FALSE`, then it returns $p(X > x) = 1 - p(X \leq x)$.

Let's look at an extreme example which is the one I mentioned above. What is the probability that $p(X < 10000)$ for $N(0, 1)$. It is almost certainly that it should be 1. In another word, $p(x > 10000)$ is 0. You can imagine the chance of having a human being whose height is 40m (ultraman).

```
## [1] 0.5
```

```
## [1] 1
```

Normal quantile

One way of defining `qnorm` is that it is the inverse of `pnorm`. So in expression `qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)`, the parameter `p` inside the `qnorm` need to be within $[0, 1]$ ($p \in [0, 1]$).

So,

```
## [1] 3.09023230616781
```

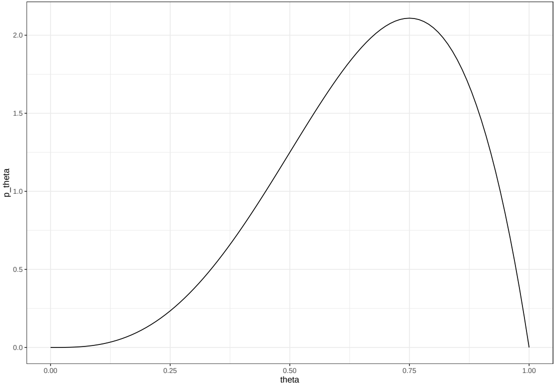
And,

```
## [1] 0.998999992234898
```

Inferring a binomial distribution

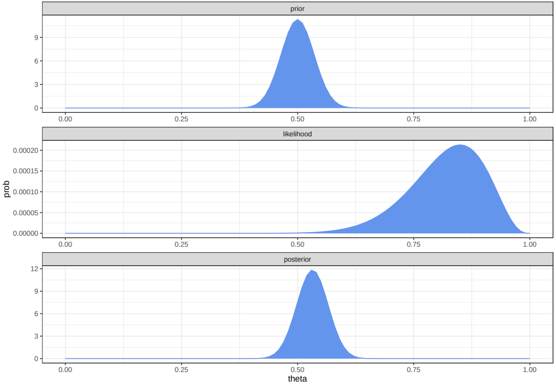
Beta Distribution

Defined here is a function `bern_beta()` which assumes a beta prior with the prior parameters, a and b to be input along with the data parameters N and z. The ouput is a faceted plot of the prior, likelihood, and posterior distributions.



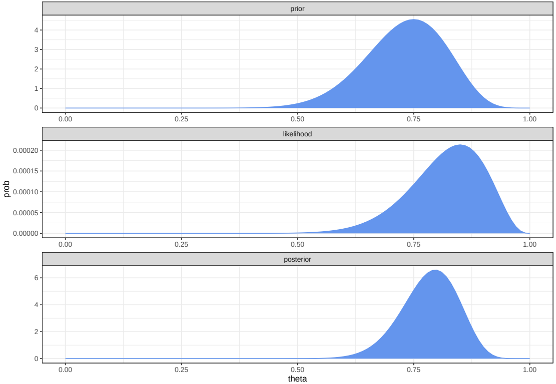
```
bern_beta(100, 100, 20, 17)

## posterior beta parameters 117 103
```



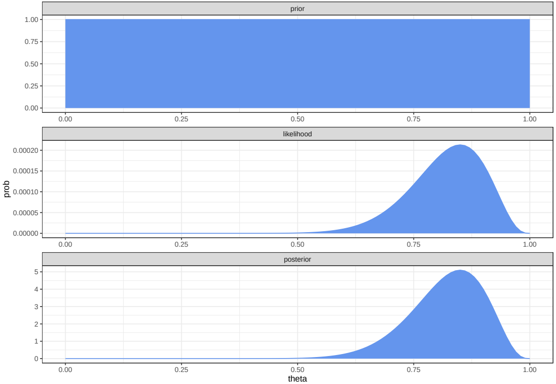
```
bern_beta(18.25, 6.75, 20, 17)

## posterior beta parameters 35.25 9.75
```



```
bern_beta(1, 1, 20, 17)
```

```
## posterior beta parameters 18 4
```



Bayesian probabilistic inference

Which coin

Create three coins, one fair and two biased:

Randomly select one of them but keep its identity secret:

Flip it 5 times and report the number of heads:

```
## [1] 3
```

Use this information to calculate likelihoods:

Update your belief about which coin is being flipped:

```
## # A tibble: 3 x 5
##   model prior likelihood product posterior
##   <fct> <dbl>      <dbl>   <dbl>     <dbl>
## 1 coin1 0.333      0.226  0.0754     0.255
## 2 coin2 0.333      0.315  0.105      0.356
## 3 coin3 0.333      0.345  0.115      0.389
```

Repeat.

Bibliography

Further study

Also see: Gupta (2002)

References

Gupta, VK, SC; Kapoor. 2002. *Fundamentals of Mathematical Statistics: A Modern Approach*. Eleventh. Sultan Chand; Sons.