Probability Formula Sheet

Autumn 2023

We have made every attempt to proof read for typographical errors, however, we cannot bear the sole responsibility for this. Any errors should be noted by the student and we will correct them promptly.

1. **Discrete Distributions**

a. PMF (§ 5.1): probability of observing a specific value x

$$f(x) = P(X = x)$$

b. CDF (§ 5.2): accumulated probability up till a specific value x

$$F(x) = P(X \le x)$$

....

c. Mean and variance (§ 7)

i. Mean: a number which represents the **average** value of random variable across separate replications of the experiement

Definition	$\mu = E[X] = \sum_{-\infty}^{\infty} x \cdot f(x)$
Linearity of Expectation	$E[aX + b] = aE[X] + b$
Law of the Unconscious Probabilist	$E[t(X)] = \sum_{-\infty}^{\infty} t(x) \cdot f(x).$

ii. Variance: a positive number which describes spread of the values of the random variable from the mean.

Definition	$\sigma^2 = Var[X] = \sum_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)$
Short cut for calculation	$\sigma^2 = E\left[X^2\right] - \mu^2.$
Variance of linear transformation	$Var[aX + b] = a^2 \cdot Var[X]$

iii. Standard deviation: positive square root of variance which is on the same units as data. It is interpretable as the *typical* deviation of the values from the mean.

iv. Chebychevs's inequality: a useful inequality which provides an upper bound for the probability that a random variable can be more than k standard deviations from the mean.

$$P(|X - \mu| \ge k\sigma) \le \frac{1}{k^2}.$$

....

d. Binomial Random Variable (§ 6)

i. A binomial random variable counts the number of successes in n independent trials where each trial results in a success (or failure) with probability π (or $1 - \pi$). We write $X \sim Binom(n, \pi)$.

ii. Binomial PMF

$$f(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}, \quad x = 0, 1, \ldots, n$$

iii. Mean of $X \sim Binom(n, \pi)$: $n\pi$

iv. Variance of $X \sim Binom(n, \pi)$: $n\pi(1-\pi)$ (proved in HW)

v. Relevant R functions:
- `dbinom(x, size, prob)` calculates $f(x) = P(X = x)$
- `pbinom(q, size, prob)` calculates $F(q) = P(X \leq q)$
- `pbinom(q, size, prob,lower.tail = F)` calculates $P(X > q)$.

. . . .

e. Geometric random variable (8.1)

i. A geometric random variable counts the number of failures *before* we see the first success when independent trials with probability π of observing a success are performed. We write $X \sim Geom(\pi)$.

ii. Geometric PMF

$$f(x) = \pi(1-\pi)^x,\ x = 0, 1, 2, \ldots$$

iii. For any non-negative integer k, we have the result

$$P(X \geq k) = (1-\pi)^k$$

iv. A geometric distribution is *memoryless*. This means for all non-negative integers x, k

$$P(X \geq x + k | X \geq k) = P(X \geq x)$$

v. Mean of $X \sim Geom(\pi)$: $\frac{1-\pi}{\pi}$.

vi. Relevant R functions:
- `dgeom(x, prob)` calculates $f(x) = P(X = x)$
- `pgeom(q, prob)` calculates $F(q) = P(X \leq q)$
- `pgeom(q, prob,lower.tail = F)` calculates $P(X > q)$.

. . . .

f. Poisson ($ 8.2)

i. The Poisson random variable counts the number of occurrences of an event over a fixed time period or within a space. We write $X \sim Poisson(\lambda)$ where λ denotes the rate of occurrence.

ii. The PMF of a Poisson can be derived from a $Binom(n, \pi)$ by setting $\pi = \frac{\lambda}{n}$ in the binomial PMF and letting $n \to \infty$.

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!},\quad x = 0, 1, 2\ldots$$

iii. Mean of $X \sim Pois(\lambda)$: λ

iv. Variance of $X \sim Pois(\lambda)$: λ

v. Relevant R functions:
- `dpois(x, lambda)` calculates $f(x) = P(X = x)$
- `ppois(q, lambda)` calculates $F(q) = P(X \leq q)$
- `ppois(q, lambda,lower.tail = F)` calculates $P(X > q)$.

---

2. **Continuous Distributions**

a. PDF and CDF (§ 9)

i. The PDF is any function which satisfies two properties:

$$f(x) \geq 0\ \forall\ x, \quad \int_{-\infty}^{\infty} f(x)dx = 1.$$

ii. Probabilities are calculated as areas under the PDF:

$$P(a \leq X < b) = \int_a^b f(x)dx.$$

Since a single value has no area, $P(X = x) = 0$ for any x however.

iii. The CDF $F(x)$ is again the accumulated probability uptil a value x:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

iv. By the Fundamental Theorem of Calculus, we can write

$$f(x) = \frac{d}{dx}F(x).$$

....

b. Mean and variance and higher moments (§ 12)

i. Mean: $\mu = E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx$

ii. Variance: $\sigma^2 = Var[X] = E\left[(X-\mu)^2\right]$

iii. The results stated in 1c. earlier holds in the continuous case as well.

....

c. Uniform random variable (§ 10)

i. The uniform random variable is the continuous analog of the equally likely model in a discrete sample space. We write $X \sim Unif(a, b)$.

ii. PDF of a uniform

$$f(x) = \frac{1}{b-a}, \quad a \leq x < b$$

iii. Mean of $X \sim Unif(a, b)$: $(a + b)/2$

iv. Variance of $X \sim Unif(a, b)$: $(b - a)^2/12$

v. The $100p$ percentile is the number q such that $P(X < q) = p$

vi. Relevant R functions:

- `dunif(x, min, max)` calculates PDF $f(x)$
- `punif(q, min, max)` calculates $F(q) = P(X \leq q)$
- `punif(q, min, max, lower.tail = F)` calculates $P(X > q)$.
- `qunif(p, min, max)` calculates the $100p$th percentile

....

d. Exponential random variable (§ 11)

i. The exponential distribution arises as the inter-event time in a Poisson model. However, it can be used as a model for any non-negative random variable! We write $X \sim Exp(\lambda)$ where $\lambda(> 0)$ is called the *rate* parameter.

ii. PDF of an exponential random variable:

$$f(x) = \lambda e^{-\lambda x}, \quad 0 \leq x < \infty$$

iii. CDF of an exponential random variable:

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & 0 \leq x \end{cases} \end{aligned}$$

iv. The exponential distribution is *memoryless*: this means for $x > k > 0$ we have the result:

$$P(X \geq x + k | X \geq k) = P(X \geq x)$$

v. Mean of $X \sim Exp(\lambda)$: $\frac{1}{\lambda}$

vi. Variance of $X \sim Exp(\lambda)$: $\frac{1}{\lambda^2}$

vii. Relevant R functions:

- `dexp(x, rate)` calculates PDF $f(x)$
- `pexp(q, rate)` calculates $F(q) = P(X \leq q)$
- `pexp(q, rate, lower.tail = F)` calculates $P(X > q)$
- `qexp(p, rate)` calculates the 100pth percentile

....

e. Normal random variable (§ 13.1)
   i. The normal distribution is often used as a model for biological measurements such as height, weight etc. It is also the limiting distribution for other models, such as the binomial, Poisson, etc. We write $X \sim Norm(\mu, \sigma)$.
   ii. We can write $X = \mu + \sigma\, Z$ where $Z \sim Norm(0, 1)$.
   iii. PDF of a normal:
   $$f(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \qquad -\infty < x < \infty$$
   iv. The mean of $X \sim Norm(\mu, \sigma)$: μ
   v. The variance of $X \sim Norm(\mu, \sigma)$: σ^2.
   vi. The 68-95-99.7 rule states that regardless of the value of μ and σ the area within 1/2/3 standard deviations of the mean is 68%/95%/99.7%.
   vii. Relevant R functions:
      - `dnorm(x, mean, sd)` calculates PDF $f(x)$
      - `pnorm(q, mean, sd)` calculates $F(q) = P(X \leq q)$
      - `pnorm(q, mean, sd, lower.tail = F)` calculates $P(X > q)$.
      - `qnorm(p, mean, sd)` calculates the 100pth percentile

Sums and Series

Binomial Theorem For any real numbers a and b and integer $n > 0$

$$(a+b)^n \quad = \quad \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

Geometric Series For any real numbers a and r ($|r| < 1$)

$$a + ar + ar^2 + ar^3 + \cdots = \frac{a}{1-r}$$

Taylor series for e^x:

$$e^x \quad = \quad 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$$