Variance, Moments, Function of a r.v.

MATH/STAT 394: Probability I
Summer 2021 A Term

Introduction to Probability
D. Anderson, T.Seppäläinen, B. Valkó

§ 3.4, 5.2

Aaron Osgood-Zimmerman

Department of Statistics

Practice solution

Practice

Consider that the time to wait your bus is modeled by $X \sim \text{Exp}(\lambda)$.

1. What is the 90/100th quantile for $\lambda = 1$?
2. Let's say that the 90/100th quantile is 20min (λ is unknown). Give an upper bound on the prob. that you wait more 30min

Solution

1.
   - Recall that the c.d.f. of $X \sim \text{Exp}(\lambda)$ is $F(t) = 1 - e^{-\lambda t}$
   - To find the pth quantile, it suffices to inverse F.
   - Indeed, denote F^{-1} the inverse of F, then
   $$\mathbb{P}(X \leq F^{-1}(p)) = F(F^{-1}(p)) = p$$
   So F^{-1} gives us the quantile. Here $p = 1 - e^{-\lambda t} \iff t = -\frac{\log(1-p)}{\lambda}$.
   - Therefore $x_{90/100} \approx 2.30$ for $\lambda = 1$
2. If we know that the 90/100th quantile is 20min, then
$$\mathbb{P}(X \geq 30) \leq \mathbb{P}(X \geq 20) = 1 - 90/100 = 10/100$$

Recap

Median Quantiles

- The median of a r.v. X is a value m s.t.

$$\mathbb{P}(X \geq m) \geq 1/2 \qquad \mathbb{P}(X \leq m) \geq 1/2$$

- The pth quantile of a r.v. X is a value x_p s.t.

$$\mathbb{P}(X \leq x_p) \geq p \qquad \mathbb{P}(X \geq x_p) \geq 1 - p$$

Variance

- Variance

$$\mathrm{Var}(X) = \mathbb{E}\left((X - \mathbb{E}[X])^2\right) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

- Standard deviation: $\sqrt{\mathrm{Var}(X)}$
- for $X \sim \mathrm{Ber}(p)$, $\mathrm{Var}(X) = p(1-p)$

Outline

Variance of a uniform dist.

Exercise

Let $X \sim \text{Unif}([a, b])$ with $a < b$, what is $\text{Var}(X)$?

Solution

- Use the formulation $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ (often more easy to compute)
- We have that $\mathbb{E}[X] = \frac{a+b}{2}$ and

$$\mathbb{E}[X^2] = \int_{-\infty}^{+\infty} x^2 f(x) dx = \frac{1}{b-a} \int_a^b x^2 dx = \frac{b^3 - a^3}{3(b-a)} = \frac{1}{3}(b^2 + ba + a^2)$$

- Therefore

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{3}(b^2 + ba + a^2) - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

- Note that as $b - a$ increases, the variance naturally increases

Other variances

Lemma
If $X \sim \text{Bin}(n, p)$, $\text{Var}(X) = np(1-p)$

Note that it is simply n times the variance of a Bernoulli.[1]

Proof Idea:

1. Use the formula $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
   $\mathbb{E}[X]$ is known so it remains to compute $\mathbb{E}[X^2]$
2. Formulate the computation of $\mathbb{E}[X^2]$ from the p.m.f.
3. Make appropriate change of var.
4. Use binomial theorem to simplify some parts

Full proof given in supp. slides

Try by yourself before looking at the solution!

[1] This can be properly justified (see MATH/STAT395)

Other variances

Lemma
If $X \sim \text{Geom}(p)$, $\text{Var}(X) = (1-p)/p^2$

Proof **Idea:**

1. Use the formula $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
   $\mathbb{E}[X]$ is known so it remains to compute $\mathbb{E}[X^2]$
2. Try to compute $\mathbb{E}[X(X-1)]$, the variance is then given by using that $\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X]$
3. Formulate $\mathbb{E}[X(X-1)]$ as computing the second derivative of $h : t \to \sum_{t=0}^{+\infty} t^k$ for $|t| < 1$
4. Use that $h(t) = 1/1 - t$ to deduce the value of the second derivative, and therefore the variance

Full proof given in supp. slides

Try by yourself before looking at the solution!

Other variances

Lemma
If $X \sim \mathrm{Exp}(\lambda)$, $\mathrm{Var}(X) = 1/\lambda^2$

Proof **Idea:**

1. Use the formula $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
   $\mathbb{E}[X]$ is known so it remains to compute $\mathbb{E}[X^2]$
2. Compute $\mathbb{E}[X^2]$ by integration by part

Full proof given in supp. slides

Try by yourself before looking at the solution!

Other variances

Classical variances

- for $X \sim \text{Ber}(p)$, $\text{Var}(X) = p(1-p)$
- for $X \sim \text{Bin}(n, p)$, $\text{Var}(X) = np(1-p)$
- for $X \sim \text{Geom}(p)$, $\text{Var}(X) = (1-p)/p^2$
- for $X \sim \text{Unif}([a, b])$, $\text{Var}(X) = (b-a)^2/12$
- for $X \sim \text{Exp}(\lambda)$, $\text{Var}(X) = 1/\lambda^2$

Note

- You should be able to easily compute variances of random variables from the definition of the variance and the p.m.f./p.d.f.
- These are mostly calculus exercises
- Beyond these exercises, the variance is a key property that will allow us to charc. r.v.

Variance properties

Motivation

- The variance is **not linear**! Instead we have the following prop.

Lemma
For a r.v. X and $a, b \in \mathbb{R}$,

$$\mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}(X)$$

Proof

$$\begin{aligned}\mathrm{Var}(aX + b) &= \mathbb{E}[(aX + b - \mathbb{E}[aX + b])^2] \\ &= \mathbb{E}[(aX + b - a\mathbb{E}[X] - b)^2] = \mathbb{E}[a^2(X - \mathbb{E}[X])] = a^2 \,\mathrm{Var}(X)\end{aligned}$$

Takeaway:

- Adding a constant to the r.v. does not change the variance
- Multiplying a r.v. by a constant a you get a *standard deviation* multiplied by a
- Remember the standard deviation is $\sqrt{\mathrm{Var}(X)}$
- Note that the standard deviation has the same 'unit' as the r.v. or the mean, while the variance has the squared of this unit

Null variance

Motivation

- The following theorem formalizes the intuition that if a r.v. does not vary (i.e. $\mathrm{Var}(X) = 0$) then it must be a constant

Theorem
For a r.v. X, $\mathrm{Var}(X) = 0$ *if and only if* $\mathbb{P}(X = a) = 1$ *for some constant* $a \in \mathbb{R}$

Proof

- Clearly if $\mathbb{P}(X = a) = 1$ then $\mathbb{E}[X] = a$ and $\mathrm{Var}(X) = 0$
- On the other hand, we consider simply the discrete case, then
$$0 = \mathrm{Var}(X) = \sum_k (k - \mu)^2 \mathbb{P}(X = k)$$
where $\mu = \mathbb{E}[X]$
- This is only possible if all the terms are zero, that is $(k - \mu)^2 \mathbb{P}(X = k) = 0$, which is equivalent to
$$k = \mu \qquad \text{or} \mathbb{P}(X = k) = 0$$
- Thus the only value k of X with $\mathbb{P}(X = k) > 0$ is $k = \mu$ and hence $\mathbb{P}(X = \mu) = 1$.

Outline

Variance

Moments

Function of a discrete r.v.

Variance computations

Moments

Motivation

- The expectation gives a first summary
- Taking the square of the r.v, we get the variance, and more interesting information
- Can we go on like that and define key properties of the r.v. as

$$\mathbb{E}[X^k] \qquad \text{or} \qquad \mathbb{E}[(X-\mu)^k]$$

- Yes, and it is a great idea, explored in more details in MATH/STAT395

Moments

Definition

*The **nth moment** of a r.v. X is*

$$\mathbb{E}[X^n]$$

*The **nth centered moment** of a r.v. X is*

$$\mathbb{E}[(X - \mathbb{E}[X])^n]$$

Notes

- The first moment is the mean
- The second moment is the squared mean
- The second centered moment is the variance
- The third centered moment is called the skewness
  It informs us about the asymmetry of the r.v.
  For example, if a r.v. is symmetric the skewness is 0
- The fourth centered moment is called the kurtosis
  It is a measure 'tailedness' or how 'heavy' (likely) the tails of the distribution are

Outline

Distribution of a function of a r.v.

Motivation

- As already observed, numerous r.v. of interest may be expressed as $Y = g(X)$ for X a classical r.v.
- We have seen how to compute $\mathbb{E}[Y]$, i.e.,
  - $\mathbb{E}[g(X)] = \sum_{k \in \mathcal{X}} g(k)\mathbb{P}(X = k)$ for X discrete
  - $\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$ for X continuous
- Now this may not be sufficient (just having the mean may not fully characterize your dist.)
- Here we are going to see how to compute the p.m.f./p.d.f. of a function of a r.v.

Inverse of a function

Motivation

- The main idea is to be able to map back values of Y onto values of X
- Recall then the definition of an inverse

Definition

Let E, F be two sets. A function $g : E \to F$ is invertible if

$$\text{for any } y \in F, \text{ there exists a unique } x \in E \text{ s.t. } y = g(x)$$

We denote the inverse of g as g^{-1} which satisfies for any $x \in E$, $y \in F$

$$g^{-1}(g(x)) = x \qquad g(g^{-1}(y)) = y$$

Example

- The function exp has inverse log
- The function x^2 is not invertible on $\mathbb{R}$
- Any strictly increasing/decreasing function $g : \mathbb{R} \to \mathbb{R}$ is invertible

Distribution of a function of a r.v.

Simple case

- Consider g invertible and X a r.v. taking values in $\mathcal{X}$
- The r.v. $Y = g(X)$ takes values in $\{g(k) : k \in \mathcal{X}\}$
- The p.m.f. of Y is then given by

$$\mathbb{P}(Y = g(k)) = \mathbb{P}(g^{-1}(Y) = k) = \mathbb{P}(X = k)$$

- Now what if g is not invertible ?

Practice next lecture

Practice

Let X be a discrete r.v. s.t.

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for } k \in \{0, 1, \ldots\}$$

1. Check that $\sum_{k=0}^{+\infty} \mathbb{P}(X = k) = 1$
2. Compute $\mathbb{E}[X], \mathrm{Var}[X]$ for $\lambda = 1$

Reminder: The Taylor series expansion of exp is $e^x = \sum_{n=0}^{+\infty} \frac{x^k}{k!}$ for any $x \in \mathbb{R}$.

Outline

Variance of a geometric r.v.

Exercise

Consider $X \sim \text{Geom}(p)$. What is $\text{Var}(X)$?

Solution

- Idea: Use that $\frac{d}{d^2 t} \sum_{k=0}^{+\infty} t^k = \frac{d}{d^2 t} \frac{1}{1-t}$, which gives for any $0 < p < 1$

$$\sum_{k=2}^{+\infty} k(k-1)t^{k-2} = \frac{2}{(1-t)^3}$$

- Here we have that

$$\mathbb{E}[X^2] = \mathbb{E}[X] + \mathbb{E}[X(X-1)]$$

and denoting $q = 1 - p$

$$\mathbb{E}[X(X-1)] = \sum_{k=2}^{+\infty} k(k-1)q^{k-1}p = pq \sum_{k=2}^{+\infty} k(k-1)q^{k-2} = \frac{2pq}{(1-q)^3} = \frac{2p(1-p)}{p^3}$$

- Therefore

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{p} + \frac{2p(1-p)}{p^3} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Variance of a binomial r.v.

Exercise

Let $X \sim \text{Bin}(n, p)$, compute $\text{Var}(X)$.

Solution

- $$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=1}^{n} k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} (k-1) \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} + \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-2)!(n-k)!} p^k (1-p)^{n-k} + \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= n(n-1)p^2 + np
\end{aligned}$$

  For the last line, for the second term you recognize the same term as the one computed for the expectation. For the first term, make a change of variable $k-2->j$, then factorizes p^2 and $n(n-1)$ and use the binomial theorem

Variance of an exponential r.v.

Exercise

Let $X \sim \text{Exp}(\lambda)$, compute $\text{Var}(X)$

Solution

- We have (recall that $\lambda > 0$ always)

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx \\
&= \left[-x^2 e^{-\lambda x}\right]_0^{+\infty} + 2\int_0^{+\infty} x e^{-\lambda x} dx \\
&= 0 + 2\left[-\frac{x}{\lambda} e^{-\lambda x}\right]_0^{+\infty} + 2\int_0^{+\infty} \frac{1}{\lambda} e^{-\lambda x} dx \\
&= \frac{2}{\lambda^2}
\end{aligned}$$

- So

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$